GenCore version 5.1.9

ucleic - nucleic search, using sw model

on: June 21, 2006, 21:26:25 ; Search time 1110 Seconds
(without alignments)
11746.040 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ence: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ing table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ched: 5244920 seqs, 3486124231 residues

l number of hits satisfying chosen parameters: 10489840

num DB seq length: 0
mum DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : N_Geneseq_8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lt o.	Score	% Query Match	Length	DB	ID	Description
1	1870	100.0	1870	10	ADB80281	Adb80281 Human pro
2	330.8	17.7	33362	13	ABD32700	Abd32700 Human can
3	324.8	17.4	31116	11	ACN44954	Acn44954 Human gen
4	324.8	17.4	31279	14	ADZ13255	Adz13255 Human can
5	324.8	17.4	58822	9	ADA02540	Ada02540 Human TCO
6	324.8	17.4	58822	10	ADB72278	Adb72278 Human TCO
7	324.8	17.4	58822	10	ADE95788	Ade95788 Human TCO
8	323.6	17.3	24934	4	AAK78433	Aak78433 Human imm
9	323.6	17.3	24934	4	AAK71542	Aak71542 Human imm
10	323.6	17.3	162450	3	AAZ86967	Aaz86967 Retinobla
11	323.4	17.3	10663	5	ABA18961	Aba18961 Human ner
12	322.6	17.3	296405	11	ACN44538	Acn44538 Human gen
13	320.4	17.1	4957	4	AAS32799	Aas32799 Human gen
14	320.4	17.1	4961	4	AAS32798	Aas32798 Human gen
15	320.4	17.1	110000	14	AEA61160_2	Continuation (3 of
16	320.4	17.1	126001	12	ADH77123	Adh77123 Human PAZ
17	320.4	17.1	272022	12	ADQ97126	Adq97126 Human can
18	319.8	17.1	6093	4	AAS32581	Aas32581 Human gen
19	318	17.0	96649	8	ACC45150	Acc45150 Human NAC
20	316	16.9	34875	10	ADD14677	Add14677 Human src
21	316	16.9	34875	13	ADR52965	Adr52965 Drug ther
22	316	16.9	34875	14	AED17821	Aed17821 Fibrotic
23	314	16.8	260160	12	ADQ20017	Adq20017 Human sof
24	313.2	16.7	60935	13	ACN37224	Acn37224 Human per
25	313	16.7	24318	13	ABD33393	Abd33393 Human can
c 26	312.8	16.7	1396	4	AAK81870	Aak81870 Human imm
c 27	312.8	16.7	89328	6	ABL61995	Abl61995 Colon ade
28	312.8	16.7	99014	6	ABN96931	Abn96931 Gene #342
c 29	312.8	16.7	106938	13	ABD33432	Abd33432 Human can
c 30	312.8	16.7	106938	13	ADR67034	Adr67034 Human can
31	312.6	16.7	5262	4	AAK71768	Aak71768 Human imm
32	312.6	16.7	5262	4	AAK71769	Aak71769 Human imm
33	312.4	16.7	193363	14	AED89392	Aed89392 Human bre
34	312.2	16.7	60729	12	ADQ97864	Adq97864 Human can
35	312	16.7	37314	4	AAK71358	Aak71358 Human imm
36	312	16.7	121062	12	ADQ97313	Adq97313 Human can
c 37	312	16.7	172570	6	ABQ88207	Abq88207 Human ost
38	311.8	16.7	13789	8	ABZ73944	Abz73944 Secreted
39	311.8	16.7	13789	10	ADC20691	Adc20691 Human sec
40	311.8	16.7	111084	12	ADQ18808	Adq18808 Human sof
c 41	311.6	16.7	26201	6	ABK14446	Abk14446 Human HMG
c 42	311.6	16.7	26201	6	ABK14039	Abk14039 Human 3-h
43	311.6	16.7	86950	14	ADX98572	Adx98572 Human HT0
c 44	311.6	16.7	107310	14	ADZ13456	Adz13456 Human can
45	311.2	16.6	19026	10	ADC86778	Adc86778 Human GPC

ALIGNMENTS

RESULT 1
ADB80281
ID ADB80281 standard; DNA; 1870 BP.
XX
AC ADB80281;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human promoter sequence #SEQ ID 26.
XX
KW PPARgamma; peroxisome proliferator activated receptor;
KW insulin resistance; drug; promoter; human; ds.
XX
OS Homo sapiens.
XX
PN WO2003062427-A1.
XX
PD 31-JUL-2003.
XX
PF 22-JAN-2003; 2003WO-JP000546.
XX
PR 23-JAN-2002; 2002JP-00013721.
PR 03-SEP-2002; 2002JP-00257703.
XX
PA (YAMA) YAMANOUCHI PHARM CO LTD.
XX
PI Endoh H, Nakano R, Kurosaki E, Kato M, Yokota H, Inabe K;
XX
DR WPI; 2003-598753/56.
XX
PT Screening for proteins that interact with peroxisome proliferator-
PT activated receptor (PPAR) for screening for drugs that improve insulin
PT resistance.
XX
PS Example 14; Page 107-109; 116pp; Japanese.
XX
CC The invention relates to a method for screening for proteins that
CC interact with peroxisome proliferator-activated receptor-gamma
CC (PPARgamma). The method of the invention is ligand-dependent and
CC comprises using a polynucleotide encoding at least residues 204-505 of a
CC fully defined 505 amino acid sequence as the bait, and a yeast two hybrid

system using a cDNA library as prey in the presence of the PPAR ligand. Also disclosed are cells that express the PPAR protein, a method for screening for drugs for insulin resistance, and preparing a medical composition for improving insulin resistance. The method of the invention is useful for screening for drugs that improve insulin resistance. The current sequence represents a human promoter sequence.

Sequence 1870 BP; 420 A; 530 C; 459 G; 461 T; 0 U; 0 Other;

ry Match 100.0%; Score 1870; DB 10; Length 1870;
t Local Similarity 100.0%; Pred. No. 0;
ches 1870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCCTGAAATAACTGCTTCT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCCTGAAATAACTGCTTCT 60

Qy         61 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 120

Qy        121 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 180

Qy        181 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 240

Qy        241 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 300

Qy        301 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 360

Qy        361 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 420

Qy        421 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 480

Qy        481 CAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTC 540

Qy        541 TTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCC 600

Qy        601 TCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCAC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCAC 660

Qy        661 GCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720

Qy        721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780

Qy        781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840

Qy        841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900

Qy        901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960

Qy        961 GCGCCAGGCCTGAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCGCCAGGCCTGAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACG 1020

Qy       1021 CTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTAC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTAC 1080

Qy       1081 CATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAGA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAGA 1140

Qy       1141 TAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATAA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATAA 1200

Qy       1201 AGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAAG 1260

Qy       1261 TGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGGA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGGA 1320

Qy       1321 GGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGGT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGGT 1380

Qy       1381 TCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAGT 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAGT 1440

Qy       1441 CCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTGC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTGC 1500

Qy       1501 TGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCGC 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCGC 1560

Qy       1561 CTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAATC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAATC 1620

Qy       1621 CCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCCA 1680

Qy       1681 CAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACCT 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACCT 1740

Qy       1741 ACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCAG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCAG 1800

Qy       1801 GCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGGA 1860

Qy       1861 TCCTTCGTAG 1870
              ||||||||||
Db       1861 TCCTTCGTAG 1870
```

RESULT 2
ABD32700/c
ID ABD32700 standard; DNA; 33362 BP.
XX

ABD32700;

18-NOV-2004 (first entry)

Human cancer-associated genomic DNA HD14-035.

Human; ds; cancer-associated protein; gene; cytostatic; cancer; leukaemia; lymphoma; CAP.

Homo sapiens.

WO2004074320-A2.

02-SEP-2004.

17-FEB-2004; 2004WO-US004730.

14-FEB-2003; 2003US-00367094.
14-MAR-2003; 2003US-00388838.
15-APR-2003; 2003US-00417375.
13-JUN-2003; 2003US-00461862.
15-SEP-2003; 2003US-00663431.
15-DEC-2003; 2003US-00737318.

(SAGR-) SAGRES DISCOVERY INC.

Morris DW, Morris DW, Malandro MS;

WPI; 2004-652914/63.

New isolated cancer-associated polynucleotides and polypeptides useful for diagnosing, preventing or treating cancers, especially lymphoma and leukemia, or in screening for agents that modulate cancer.

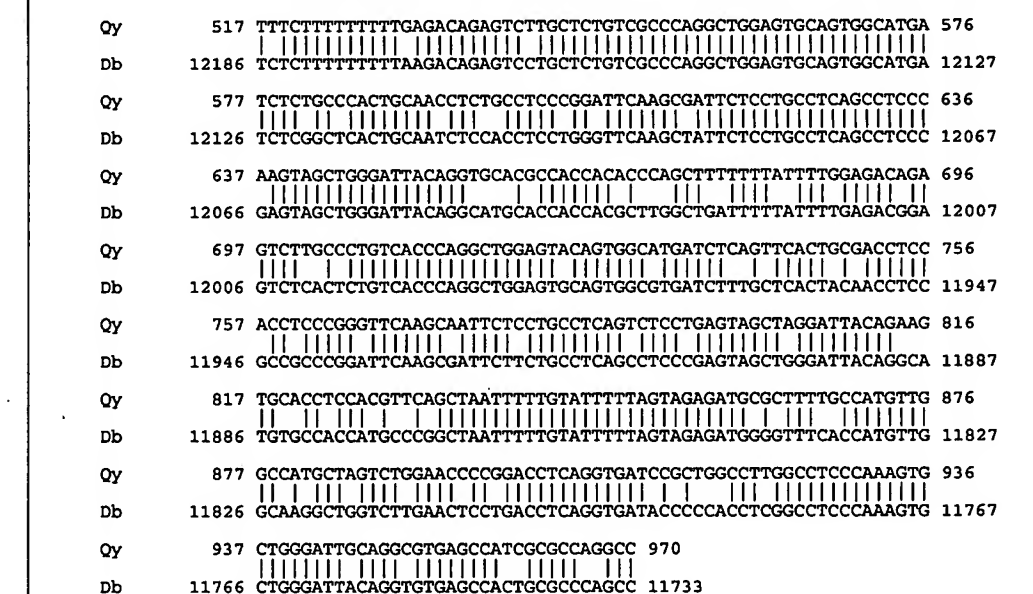
claim 16; seqid 265; 310pp; English.

The invention relates to an isolated nucleic acid comprising at least 10 contiguous nucleotides of any of the 233 polynucleotide sequences given in the specification, or its complement. The nucleic acids encode cancer-associated proteins. Also included are an expression vector comprising the isolated nucleic acid cited above, a host cell comprising the above recombinant nucleic acid or expression vector, a microarray for detecting a cancer-associated (CA) nucleic acid comprising at least one probe comprising at least 10 contiguous nucleotides of any of the above-mentioned nucleotide sequences, an isolated polypeptide (encoded within an open reading frame of a CA sequence selected from any of the 95 polynucleotide sequences as mentioned in the specification, or its complement), an isolated antibody, (or its antigen binding fragment) that binds to the above polypeptide, a hybridoma that produces the above monoclonal antibody, a pharmaceutical composition comprising the above antibody and a pharmaceutical excipient, a kit for detecting cancer cells(comprising the antibody cited above, methods for diagnosing cancer or for detecting the presence or absence of cancer cells in an individual, a method for inhibiting growth of cancer cells in an individual, a method for delivering a therapeutic agent to cancer cells in an individual, an electronic library comprising the above polynucleotide or polypeptide (or their fragments), methods of screening for anticancer activity or for a bioactive agent capable of modulating the activity of a CA protein (CAP), methods for detecting cancer associated with expression of a polypeptide in a test cell sample, a method for treating cancers and a method for inhibiting the expression of CA gene in a cell. The composition and methods are useful for detecting, diagnosing, preventing and treating cancers, especially lymphoma and leukaemia. These may also be used in screening for agents that modulate cancer. The present sequence is a human CAP genomic sequence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 33362 BP; 8042 A; 8668 C; 8695 G; 7957 T; 0 U; 0 Other;

ery Match 17.7%; Score 330.8; DB 13; Length 33362;
st Local Similarity 83.0%; Pred. No. 1.1e-74;

Matches 377; Conservative 0; Mismatches 77; Indels 0; Gaps 0;



RESULT 3
ACN44954
ID ACN44954 standard; DNA; 31116 BP.
XX
AC ACN44954;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG38622.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1660; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for

evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 31116 BP; 7214 A; 8217 C; 7722 G; 7963 T; 0 U; 0 Other;

:ery Match 17.4%; Score 324.8; DB 11; Length 31116;
:t Local Similarity 77.5%; Pred. No. 3.7e-73;
:ches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

```
Qy       502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
            ||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
Db     13037 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 13096

Qy       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
            | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
Db     13097 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 13156

Qy       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
            |||||||||||||  |||||||||| |||||||| | |||| || ||| ||
Db     13157 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 13216

Qy       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
             || |||| |||| ||||  ||| ||  | ||||  |||||||||||| || |||||||
Db     13217 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 13276

Qy       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
            ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||| ||||
Db     13277 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 13336

Qy       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
             |||||||| |||||||||  |  | || |||||| | |||||||||| ||||||||||
Db     13337 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 13396

Qy       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
            ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
Db     13397 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 13456

Qy       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
            |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
Db     13457 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 13516

Qy       975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
            | ||  ||  | | |   |||| |   |   |  | | | |||
Db     13517 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 13564
```

,T 4
1255
ADZ13255 standard; DNA; 31279 BP.

ADZ13255;

16-JUN-2005 (first entry)

Human cancer-associated genomic DNA #63.

Diagnosis; DNA microarray; microarray; biochip; cancer; neoplasm; cytostatic; gene; ds.

Homo sapiens.

WO2005031001-A2.

07-APR-2005.

23-SEP-2004; 2004WO-US031617.

XX
PR 23-SEP-2003; 2003US-00669920.
XX
PA (CHIR) CHIRON CORP.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2005-273395/28.
XX
PT Nucleic acid array useful for detecting cancer associated nucleic acid,
PT comprises two or more nucleic acid probes.
XX
PS Disclosure; SEQ ID NO 775; 198pp; English.
XX
CC The invention relates to a nucleic acid array for detecting a cancer
CC associated (CA) nucleic acid, comprising two or more nucleic acid probes.
CC The invention also relates to a peptide array comprising two or more
CC isolated polypeptides encoded by a CA nucleic acid sequence, a compound
CC that binds to a polypeptide, an isolated antibody or its fragment which
CC binds to a polypeptide, which is prepared by immunizing a host animal
CC with a composition comprising the polypeptide or its antigen binding
CC fragment and collecting cells from the host expressing antibodies against
CC the antigen or its antigen binding fragment, a composition comprising the
CC antibody and a carrier, a method of screening for anticancer activity, a
CC method of detecting a CA nucleic acid, a method of diagnosing cancer, a
CC method of treating cancer and a method of inhibiting expression of a CA
CC nucleic acid in a cell. The CA nucleic acids are useful for detecting CA
CC nucleic acids. The antibody is useful for detecting the presence or
CC absence of cancer cells in an individual which involves contacting cells
CC from the individual with the antibody and detecting a complex of a CA
CC protein from the cancer cells and the antibody, where the detection of
CC the complex correlates with the presence of cancer cells in the
CC individual. The composition is useful for inhibiting growth of cancer
CC cells in an individual or for delivering a therapeutic agent to cancer
CC cells in an individual. The invention is also useful for diagnosing
CC cancer, for treating cancer and for inhibiting expression of a CA gene in
CC a cell. This sequence represents human cancer-associated genomic DNA of
CC the invention.
XX
SQ Sequence 31279 BP; 7246 A; 8268 C; 7755 G; 8010 T; 0 U; 0 Other;

Query Match 17.4%; Score 324.8; DB 14; Length 31279;
Best Local Similarity 77.5%; Pred. No. 3.7e-73;
Matches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

```
Qy       502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
            ||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
Db     13193 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 13252

Qy       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
            | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
Db     13253 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 13312

Qy       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
            |||||||||||||  |||||||||| |||||||| | |||| || ||| ||
Db     13313 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 13372

Qy       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
             || |||| |||| ||||  ||| ||  | ||||  |||||||||||| || |||||||
Db     13373 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 13432

Qy       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
            ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||| ||||
Db     13433 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 13492

Qy       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
             |||||||| |||||||||  |  | || |||||| | |||||||||| ||||||||||
Db     13493 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 13552

Qy       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
            ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
Db     13553 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 13612
```

```
      915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
          |  | |||||| ||| ||||||||||||||| |||| |||||||| |||||| |||||
    13613 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 13672

      975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
          | ||   ||  | | |    |||| |   |     | | | |||
    13673 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 13720
```

RESULT 5
ADA02540/c
ADA02540 standard; DNA; 58822 BP.

ADA02540;

06-NOV-2003 (first entry)

Human TCOF1 carcinoma associated gene, SEQ ID NO:1058.

Human; carcinoma associated; oncogene; carcinoma; cancer; breast; prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening; gene; ds.

Homo sapiens.

WO2003057146-A2.

17-JUL-2003.

26-DEC-2002; 2002WO-US041414.

26-DEC-2001; 2001US-00035832.

(SAGR-) SAGRES DISCOVERY.

Morris DW;

WPI; 2003-587068/55.

New recombinant nucleic acid encoding carcinoma associated protein, useful for preparing compositions for treating carcinomas.

Claim 1; SEQ ID NO 1058; 245pp; English.

The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 58822 BP; 14199 A; 14875 C; 15625 G; 13656 T; 0 U; 467 Other;

ery Match 17.4%; Score 324.8; DB 9; Length 58822;
t Local Similarity 77.5%; Pred. No. 4.8e-73;
tches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

```
Qy       502 TTTTTTTTTTTTCTTTTCTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
            |||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
Db     58246 TTTTTTTTTTTTTTTTGTTATGTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 58187

Qy       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
            | |||||||| |||||| || ||| ||||||||||||||||| | |||||| |||||||
Db     58186 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 58127

Qy       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
            |||||||||||||| |||||||||| ||||||||| | |||| || ||| ||
Db     58126 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 58067

Qy       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
             || |||| |||| |||  ||| ||  | ||||  |||||||||||| || |||||||
Db     58066 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 58007

Qy       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
            ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||| ||||
Db     58006 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 57947

Qy       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
            |||||||| |||||||||  | || |||||| | |||||||||||| ||||||||||
Db     57946 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 57887

Qy       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
            ||||  | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
Db     57886 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 57827

Qy       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
            |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
Db     57826 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 57767

Qy       975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
            | ||   ||  | | |    |||| |   |     | | | |||
Db     57766 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 57719
```

RESULT 6
ADB72278/c
ID ADB72278 standard; DNA; 58822 BP.
XX
AC ADB72278;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human TCOF1 gene.
XX
KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Homo sapiens.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX

Claim 1; SEQ ID NO 106; 2304pp; English.

The invention relates to a novel recombinant nucleic acid comprising a nucleotide sequence selected from any of the 660 sequences fully defined in the specification. A polynucleotide of the invention has cytostatic activity, and may have a use in gene therapy, or in a vaccine. The recombinant nucleic acids and polypeptides are useful for treating carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and sarcomas. The present sequence represents a human gene of the invention.

Sequence 58822 BP; 14199 A; 14875 C; 15625 G; 13656 T; 0 U; 467 Other;

ery Match 17.4%; Score 324.8; DB 10; Length 58822;
st Local Similarity 77.5%; Pred. No. 4.8e-73;
ches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

```
     502 TTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
         ||||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
   58246 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 58187

     562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
         | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
   58186 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 58127

     622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
         ||||||||||||||  |||||||||| |||||||| | |||| || ||| ||
   58126 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 58067

     676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
          || |||| |||| ||||  ||| ||  | ||||  |||||||||||||| || |||||||
   58066 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 58007

     735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
         ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||||| ||||
   58006 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 57947

     795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
         ||||||||  |||||||||  |  | || ||||||  | ||||||||||  ||||||||||
   57946 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 57887

     855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
         ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||| ||||
   57886 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 57827

     915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
         |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
   57826 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 57767

     975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
         | ||   ||  | | |    |||| |   |     |  | | | |||
   57766 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 57719
```

RT 7
5788/c
ADE95788 standard; DNA; 58822 BP.

ADE95788;

12-FEB-2004 (first entry)

Human TCOF1 gene genomic DNA sequence.

cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy; lymphoma; breast cancer; prostate cancer; leukaemia; ds; human; TCOF1.

Homo sapiens.

WO2003039484-A2.

15-MAY-2003.

08-NOV-2002; 2002WO-US036071.

XX
PR 08-NOV-2001; 2001US-00052482.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-441462/41.
XX
PT New carcinoma associated nucleic acids and proteins, useful for screening
PT drug candidates, or for diagnosing and treating carcinomas, e.g.
PT lymphoma, breast cancer, prostate cancer or leukemia.
XX
PS Claim 1; SEQ ID NO 46; 793pp; English.
XX
CC This invention relates to novel recombinant nucleic acids for use in
CC diagnosis and treatment of cancer, especially carcinomas, as well as the
CC use of compositions in screening methods. The compositions of the
CC invention may have cytostatic activity whilst the disclosed sequences may
CC be useful for gene therapy. The carcinoma associated nucleic acids and
CC proteins are useful for diagnosing and treating carcinomas, for example
CC lymphoma, breast cancer, prostate cancer or leukaemia, or for screening
CC drug candidates or bioactive agents capable of binding to, or modulating
CC the activity of, a carcinoma associated protein. The present sequence is
CC the genomic DNA sequence of the human TCOF1 gene which is a carcinoma
CC associated gene of the invention.
XX
SQ Sequence 58822 BP; 14199 A; 14881 C; 15619 G; 13656 T; 0 U; 467 Other;

Query Match 17.4%; Score 324.8; DB 10; Length 58822;
Best Local Similarity 77.5%; Pred. No. 4.8e-73;
Matches 409; Conservative 0; Mismatches 112; Indels 7; Gaps 1;

```
Qy       502 TTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
             ||||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
Db     58246 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 58187

Qy       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
             | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
Db     58186 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 58127

Qy       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
             ||||||||||||||  |||||||||| |||||||| | |||| || ||| ||
Db     58126 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 58067

Qy       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
              || |||| |||| ||||  ||| ||  | ||||  |||||||||||||| || |||||||
Db     58066 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 58007

Qy       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
             ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||||| ||||
Db     58006 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 57947

Qy       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
             ||||||||  |||||||||  |  | || ||||||  | ||||||||||  ||||||||||
Db     57946 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 57887

Qy       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
             ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||| ||||
Db     57886 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 57827

Qy       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
             |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
Db     57826 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 57767

Qy       975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
             | ||   ||  | | |    |||| |   |     |  | | | |||
Db     57766 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 57719
```

RESULT 8
AAK78433/c
ID AAK78433 standard; DNA; 24934 BP.

AAK78433;

07-NOV-2001 (first entry)

Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33245.

Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.

Homo sapiens.

WO200157182-A2.

09-AUG-2001.

17-JAN-2001; 2001WO-US001354.

31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.
16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198123P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.

06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 33245; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention

Sequence 24934 BP; 7325 A; 4578 C; 4828 G; 8203 T; 0 U; 0 Other;

ery Match 17.3%; Score 323.6; DB 4; Length 24934;
st Local Similarity 81.2%; Pred. No. 7e-73;
tches 401; Conservative 0; Mismatches 89; Indels 4; Gaps 2;

```
      488 TGAGCTGCTGCTGCTTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
          | ||  || |  |   ||.|||||| ||||   || ||||||||||||||||||||||||
    18989 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 18930

      548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
          |||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
    18929 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 18870

      608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
          ||||||| || || |||||||||||||| |||||||||||| |||||  |  |||||||
    18869 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 18810

      668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
          | ||||||||  ||||||  ||||  |||  |||||| || ||||||||||||||||||| |
    18809 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 18750

      726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
          |||||| ||||||| | ||||||| | |||| |||||||||||  |  | ||||||||||
    18749 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 18690

      786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
          |||| |||||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
    18689 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 18630

      844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
          | ||||||||||||||  | |||  ||||||| |||| | | |||| || | || |||||
    18629 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 18570

      904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
          |||||||||| ||  |||||||||||||||||||||||||||||| |||||||||||| ||||
Db    18569 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 18510

Qy      964 CCAGGCCTGAGCTA 977
            ||  |||  | | |
Db    18509 CCCAGCCACACCCA 18496
```

RESULT 9
AAK71542/c
ID AAK71542 standard; DNA; 24934 BP.
XX
AC AAK71542;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26354.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.

08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
25-SEP-2000; 2000US-0234998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.
08-NOV-2000; 2000US-0246610P.
08-NOV-2000; 2000US-0246611P.
08-NOV-2000; 2000US-0246613P.
17-NOV-2000; 2000US-0249207P.
17-NOV-2000; 2000US-0249208P.
17-NOV-2000; 2000US-0249209P.
17-NOV-2000; 2000US-0249210P.
17-NOV-2000; 2000US-0249211P.
17-NOV-2000; 2000US-0249212P.
17-NOV-2000; 2000US-0249213P.
17-NOV-2000; 2000US-0249214P.
17-NOV-2000; 2000US-0249215P.
17-NOV-2000; 2000US-0249216P.
17-NOV-2000; 2000US-0249217P.
17-NOV-2000; 2000US-0249218P.
17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 26354; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 24934 BP; 7325 A; 4578 C; 4828 G; 8203 T; 0 U; 0 Other;

Query Match 17.3%; Score 323.6; DB 4; Length 24934;
Best Local Similarity 81.2%; Pred. No. 7e-73;
Matches 401; Conservative 0; Mismatches 89; Indels 4; Gaps 2;

```
Qy      488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
            | ||  || |  |   || ||||||| ||||   || ||||||||||||||||||||||
Db    18989 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 18930

Qy      548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
            |||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
Db    18929 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 18870

Qy      608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
            ||||||| || || |||||||||||||||| |||||||||||| |||||  |  |||||||
Db    18869 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 18810

Qy      668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
            | |||||||  ||||||||||  |||  |||||| || ||||||||||||||||||| |
Db    18809 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 18750

Qy      726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
            |||||| ||||||| | ||||||| | |||| |||||||||||  |  ||||||||||
```

```
  18749 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 18690

    786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
        |||| |||||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
  18689 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 18630

    844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
        | |||||||||||||    | |||  ||||||| |||| | | |||| || | || |||||
  18629 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 18570

    904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
        ||||||||| ||  |||||||||||||||||||||||||||| |||||||||||| ||||
  18569 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 18510

    964 CCAGGCCTGAGCTA 977
        ||  |||  | | |
  18509 CCCAGCCACACCCA 18496
```

```
;T 10
;967/c
AAZ86967 standard; DNA; 162450 BP.

AAZ86967;

16-MAY-2000  (first entry)

Retinoblastoma binding protein-7 genomic DNA sequence.

RBP-7; retinoblastoma binding protein-7; abnormal cell proliferation;
diagnosis; therapy; cell differentiation; thyroid hyperplasia; psoriasis;
benign prostate hypertrophy; cancer; sarcoma; neoplasm; leukaemia;
lymphoma; ds.

Homo sapiens.

WO200000607-A1.

06-JAN-2000.

30-JUN-1999;   99WO-IB001242.

30-JUN-1998;   98US-0091315P.
10-DEC-1998;   98US-0111909P.

(GEST ) GENSET.

Bougueleret L;

WPI; 2000-117170/10.

Novel nucleic acid and polymorphic markers used for diagnosis of
diseases, especially those involving abnormal cell proliferation and
differentiation.

Claim 1; Page 118-163; 223pp; English.

This sequence represents the retinoblastoma binding protein-7 (RBP-7)
genomic sequence of the invention. The RBP-7 coding sequence and
regulatory sequences are useful for the recombinant production of the
protein and for expressing heterologous nucleic acids. Primers and probes
derived from the RBP-7 nucleotide sequence (e.g. AAZ87035-Z87099) are
useful for DNA amplification and detection methods. RBP-7 biallelic
markers (see AAZ86993-Z87034) are useful for diagnosis of disease related
to alteration in the regulation or in the coding regions of the RBP-7
gene and for prognosis/diagnosis of an eventual treatment with
therapeutic agents, especially agents acting on pathologies involving
abnormal cell proliferation and/or differentiation, these include thyroid
hyperplasia, psoriasis, benign prostate hypertrophy, cancers, including
breast cancer, sarcomas and other neoplasms, bladder cancer, colon
cancer, lung cancer, prostate cancer, various leukaemias, and lymphomas.
RBP-7 antibodies are useful as diagnostic agents
```

```
SQ   Sequence 162450 BP; 45465 A; 30661 C; 32637 G; 53673 T; 0 U; 14 Other;

  Query Match             17.3%;  Score 323.6;  DB 3;  Length 162450;
  Best Local Similarity   81.2%;  Pred. No. 1.5e-72;
  Matches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              | ||  || | |   || ||||||| ||||   || ||||||||||||||||||||||
Db     102008 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
              |||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
Db     101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

Qy        608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
              ||||||| || || |||||||||||||| |||||||||||| |||||  | |||||||
Db     101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

Qy        668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
              | |||||||  |||||| ||||  |||  |||||| || ||||||||||||||||||| |
Db     101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

Qy        726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
              |||||| ||||||| | ||||||| | |||| |||||||||| |  | ||||||||||
Db     101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709

Qy        786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
              |||| |||||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
Db     101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

Qy        844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
              | |||||||||||||    | |||  ||||||| |||| | | |||| || | || |||||
Db     101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

Qy        904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
              ||||||||| ||  |||||||||||||||||||||||||||| |||||||||||| ||||
Db     101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

Qy        964 CCAGGCCTGAGCTA 977
              ||  |||  | | |
Db     101528 CCCAGCCACACCCA 101515
```

```
RESULT 11
ABA18961
ID   ABA18961 standard; DNA; 10663 BP.
XX
AC   ABA18961;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 11292.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001334.
XX
PR   31-JAN-2000; 2000US-0179065P.
PR   04-FEB-2000; 2000US-0180628P.
PR   24-FEB-2000; 2000US-0184664P.
PR   02-MAR-2000; 2000US-0186350P.
```

16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198123P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
25-SEP-2000; 2000US-0234998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
PS Disclosure; SEQ ID NO 11292; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins

(ABB14678-ABB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 10663 BP; 2921 A; 2087 C; 2099 G; 3555 T; 0 U; 1 Other;

ıry Match 17.3%; Score 323.4; DB 5; Length 10663;
ıt Local Similarity 81.3%; Pred. No. 5.6e-73;
ıches 391; Conservative 0; Mismatches 81; Indels 9; Gaps 1;

```
  502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
      |||||||||||| ||||| |||||||||||||  |||||   |||||| ||||||||||
 5384 TTTTTTTTTTTTTTTTTTTTTTTTTTTGAGATGGAGTCACACTCTGTTGCCCAGGCTGG 5443

  562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
      ||||  ||||    ||||  |  ||||||||||||||||||| | | |||||| |||||||
 5444 AGTGTCGTGGTGCAATCTTGGGTCACTGCAACCTCTGCCTCACAGGTTCAAGTGATTCTC 5503

  622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
      |||||||||||||  |||||||  | | ||||||  |  ||||||||  ||||||
 5504 CTGCCTCAGCCTCCTGAGTAGCTAGTACTACAGGCACGTGCCACCACGCCCAGCTAATTT 5563

  676 ---TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCA 732
         ||||||| |||| ||||||||||||  | |||||  |||||||||||| |||| |||
 5564 TTGTTTTTTTGTTTTTGAGACAGAGTCTCACTCTGTCGGCCAGGCTGGAGTGCAGTAGCA 5623

  733 TGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCT 792
      ||||||||  |||||||  |||||| |||| |||| |||| ||||||||||||||||| ||
 5624 CGATCTCAGCTCACTGCAACCTCCGCCTCGCGGGCTCAAACAATTCTCCTGCCTCAGCCT 5683

  793 CCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAG 852
      ||  ||||||| |||||||||  |||  || ||| |  | ||||||||||||||||||||
 5684 CCCAAGTAGCTGGGATTACAGGTGTGTGCCACCATGCCCGGCTAATTTTTGTATTTTTAG 5743

  853 TAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATC 912
      ||||||  |  |||  |||||||||||| ||| |||| |||| || |||| |||||  |||
 5744 TAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCACAGGTAATC 5803

  913 CGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTG 972
      |||| |||| |||||||||||||||||||||| |||||||||||| ||  || |||||
 5804 CGCTCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTTCCCGGCCTA 5863

  973 A 973
      |
 5864 A 5864
```

ıT 12
ı538
ACN44538 standard; DNA; 296405 BP.

ACN44538;

18-NOV-2004 (first entry)

Human genomic sequence hCG16326.

Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

Homo sapiens.

XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1036; 0pp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
SQ Sequence 296405 BP; 82162 A; 54238 C; 57164 G; 90624 T; 0 U; 12217 Other;

Query Match 17.3%; Score 322.6; DB 11; Length 296405;
Best Local Similarity 79.7%; Pred. No. 3.5e-72;
Matches 396; Conservative 0; Mismatches 94; Indels 7; Gaps 1;

```
Qy       488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
             || | || |||    |   | |||||  |||  || |||||||  ||  ||||||  ||||
Db    117715 TGTGATGATGCAATATGAATCTTTTTTGTTTGTTTATTTTTTGGGATGGAGTCTCACTCT 117774

Qy       548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
             ||  |||||||||||||||||||| |  ||||| || | |||||||||||||||||| |
Db    117775 GTAACCCAGGCTGGAGTGCAGTGGTACAATCTCAGCTCGCTGCAACCTCTGCCTCCCCGG 117834

Qy       608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
             ||| ||| |||||||||||||||||||||||||||||||||||||||||  | |||||||
Db    117835 TTCCAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCACCTGCCACCA 117894

Qy       668 CACCCAGCT-------TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720
               | |||||       ||| ||  ||| |||||||||||||| | ||||  ||||||||||
Db    117895 TGCTCAGCTAATTTTGTTTGTTTGTTTTGAGACAGAGTCTTACTCTGTGGCCCAGGCTGG 117954

Qy       721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780
             ||| ||||||||  ||||| | || |||| |||||||||||||| ||||| ||| |||||
Db    117955 AGTGCAGTGGCACAATCTCGGCTCGCTGCAACCTCCACCTCCCCGGTTCCAGCTATTCTT 118014

Qy       781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840
             |||||||||||||| |||||||| |||||| ||  |||  |  |||||  | ||||||||
Db    118015 CTGCCTCAGTCTCCCGAGTAGCTGGGATTATAGGCGTGTGCTACCACGCCCGGCTAATTT 118074

Qy       841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900
             |||||||||||||||||| | | |||  |||||||||||| ||| |||| |||| || |
Db    118075 TTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGG 118134

Qy       901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960
             ||||| ||||||| |  |||||||||||||||||||||||||||||||| |||| |||||||| |
Db    118135 CCTCAAGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACC 118194
```

```
      961 GCGCCAGGCCTGAGCTA 977
            ||| |||||||  ||
   118195 CTGCCCGGCCTGAAATA 118211
```

LT 13
:799/c
AAS32799 standard; DNA; 4957 BP.

AAS32799;

17-DEC-2001 (first entry)

Human genomic DNA for novel endocrine antigen, SEQ ID No 753.

Human; endocrine antigen; ds; cytostatic; antiinfertility; antidiabetic; thyroid-active; adrenal-active; androgenic; gastric; gene therapy; antisense-therapy; antibody; endocrine disorder; hormone imbalance; reproductive disorder; endocrine cancer; pancreatic disorder; diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder; hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.

Homo sapiens.

WO200155319-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001335.

31-JAN-2000; 2000US-0179065P.
04-FEB-2000; 2000US-0180628P.
24-FEB-2000; 2000US-0184664P.
02-MAR-2000; 2000US-0186350P.
16-MAR-2000; 2000US-0189874P.
17-MAR-2000; 2000US-0190076P.
18-APR-2000; 2000US-0198123P.
19-MAY-2000; 2000US-0205515P.
07-JUN-2000; 2000US-0209467P.
28-JUN-2000; 2000US-0214886P.
30-JUN-2000; 2000US-0215135P.
07-JUL-2000; 2000US-0216647P.
07-JUL-2000; 2000US-0216880P.
11-JUL-2000; 2000US-0217487P.
11-JUL-2000; 2000US-0217496P.
14-JUL-2000; 2000US-0218290P.
26-JUL-2000; 2000US-0220963P.
26-JUL-2000; 2000US-0220964P.
14-AUG-2000; 2000US-0224518P.
14-AUG-2000; 2000US-0224519P.
14-AUG-2000; 2000US-0225213P.
14-AUG-2000; 2000US-0225214P.
14-AUG-2000; 2000US-0225266P.
14-AUG-2000; 2000US-0225267P.
14-AUG-2000; 2000US-0225268P.
14-AUG-2000; 2000US-0225270P.
14-AUG-2000; 2000US-0225447P.
14-AUG-2000; 2000US-0225757P.
14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.

17-NOV-2000; 2000US-0249244P.
17-NOV-2000; 2000US-0249245P.
17-NOV-2000; 2000US-0249264P.
17-NOV-2000; 2000US-0249265P.
17-NOV-2000; 2000US-0249297P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-457726/49.

Isolated polypeptide for treating, preventing and prognosing disorders related to the endocrine system including endocrine disorders, reproductive disorders, and gastrointestinal disorders and also for testing and detection e.g. diagnosis.

Disclosure; SEQ ID NO 753; 558pp; English.

The invention relates to cDNAs encoding novel human endocrine antigens or a fragment having biological activity, a domain, an epitope, full length protein, variant, allelic variant or a species homologue of the cDNA/antigen. The DNAs and polypeptides are useful for preventing, treating or ameliorating a medical condition when administered (e.g. by gene therapy or antisense-therapy). Identifying mutations in the genes coding for the antigens is useful for diagnosing a pathological condition or a susceptibility to a pathological condition. The DNAs, antigens and antibodies raised against the antigens useful for treating, preventing and/ or prognosing disorders related to the endocrine system or hormone imbalance or reproductive disorders, cancers of endocrine tissues, disorders of the pancreas (e.g. diabetes mellitus), the adrenal glands (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the hypothalamus and testes (e.g. vanishing testes syndrome), many examples of diseases and disorders are given in the specification. The present sequence is genomic DNA fragment form a gene encoding an endocrine antigen of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 4957 BP; 1243 A; 1249 C; 1402 G; 1063 T; 0 U; 0 Other;

ry Match 17.1%; Score 320.4; DB 4; Length 4957;
t Local Similarity 82.8%; Pred. No. 2.5e-72;
ches 394; Conservative 0; Mismatches 71; Indels 11; Gaps 2;

```
      506 TTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
          ||| |||| ||||| |||||||| ||||  |||||||||||||| ||||| |||||||||
     1768 TTTGTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCTGTAGCCCAAGCTGGAGTG 1709

      566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
          |||||||||||||||| ||| ||||||||| |  |||||||| |||||||  |||||||||
     1708 CAGTGGCATGATCTCGGCCTACTGCAACCACCACCTCCCGGGTTCAAGCAGTTCTCCTGC 1649

      626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATT 685
          |||||||||||||||||||||||||||||||  | ||||||||||||||| ||||| ||
     1648 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGCACCACCACACCCAGCTATTTTTGTT 1589

      686 TTG---------GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          |||         ||||| |||||||||  ||||| |||||||||||| |||||| |||||
Db    1588 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 1529

Qy     737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
           ||| | ||||||| |||||| |||||||||||||||  |||||||||||||| |||| |
Db    1528 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 1469

Qy     797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
           ||||||| ||  ||||| | |  |||||||     |||||  ||||||||||||||||
Db    1468 AGTAGCTGGGCCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 1409

Qy     855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
           |||| |   |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
Db    1408 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 1349

Qy     915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
           |  |||| |||||||||||||||||||||||| ||||| ||||||| |  ||| ||||
Db    1348 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCCTGCCCGGCC 1293
```

RESULT 14
AAS32798/c
ID AAS32798 standard; DNA; 4961 BP.
XX
AC AAS32798;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human genomic DNA for novel endocrine antigen, SEQ ID No 752.
XX
KW Human; endocrine antigen; ds; cytostatic; antiinfertility; antidiabetic;
KW thyroid-active; adrenal-active; androgenic; gastric; gene therapy;
KW antisense-therapy; antibody; endocrine disorder; hormone imbalance;
KW reproductive disorder; endocrine cancer; pancreatic disorder;
KW diabetes mellitus; adrenal gland disorder; hirsutism; thyroid disorder;
KW hyperthyroidism; hypothalamic disorder; vanishing testes syndrome.
XX
OS Homo sapiens.
XX
PN WO200155319-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001335.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.
14-AUG-2000; 2000US-0225759P.
18-AUG-2000; 2000US-0226279P.
22-AUG-2000; 2000US-0226681P.
22-AUG-2000; 2000US-0226868P.
22-AUG-2000; 2000US-0227182P.
23-AUG-2000; 2000US-0227009P.
30-AUG-2000; 2000US-0228924P.
01-SEP-2000; 2000US-0229287P.
01-SEP-2000; 2000US-0229343P.
01-SEP-2000; 2000US-0229344P.
01-SEP-2000; 2000US-0229345P.
05-SEP-2000; 2000US-0229509P.
05-SEP-2000; 2000US-0229513P.
06-SEP-2000; 2000US-0230437P.
06-SEP-2000; 2000US-0230438P.
08-SEP-2000; 2000US-0231242P.
08-SEP-2000; 2000US-0231243P.
08-SEP-2000; 2000US-0231244P.
08-SEP-2000; 2000US-0231413P.
08-SEP-2000; 2000US-0231414P.
08-SEP-2000; 2000US-0232080P.
08-SEP-2000; 2000US-0232081P.
12-SEP-2000; 2000US-0231968P.
14-SEP-2000; 2000US-0232397P.
14-SEP-2000; 2000US-0232398P.
14-SEP-2000; 2000US-0232399P.
14-SEP-2000; 2000US-0232400P.
14-SEP-2000; 2000US-0232401P.
14-SEP-2000; 2000US-0233063P.
14-SEP-2000; 2000US-0233064P.
14-SEP-2000; 2000US-0233065P.
21-SEP-2000; 2000US-0234223P.
21-SEP-2000; 2000US-0234274P.
25-SEP-2000; 2000US-0234997P.
25-SEP-2000; 2000US-0234998P.
26-SEP-2000; 2000US-0235484P.
27-SEP-2000; 2000US-0235834P.
27-SEP-2000; 2000US-0235836P.
29-SEP-2000; 2000US-0236327P.
29-SEP-2000; 2000US-0236367P.
29-SEP-2000; 2000US-0236368P.
29-SEP-2000; 2000US-0236369P.
29-SEP-2000; 2000US-0236370P.
02-OCT-2000; 2000US-0236802P.
02-OCT-2000; 2000US-0237037P.
02-OCT-2000; 2000US-0237038P.
02-OCT-2000; 2000US-0237039P.
02-OCT-2000; 2000US-0237040P.
13-OCT-2000; 2000US-0239935P.
13-OCT-2000; 2000US-0239937P.
20-OCT-2000; 2000US-0240960P.
20-OCT-2000; 2000US-0241221P.
20-OCT-2000; 2000US-0241785P.
20-OCT-2000; 2000US-0241786P.
20-OCT-2000; 2000US-0241787P.
20-OCT-2000; 2000US-0241808P.
20-OCT-2000; 2000US-0241809P.
20-OCT-2000; 2000US-0241826P.
01-NOV-2000; 2000US-0244617P.
08-NOV-2000; 2000US-0246474P.
08-NOV-2000; 2000US-0246475P.
08-NOV-2000; 2000US-0246476P.
08-NOV-2000; 2000US-0246477P.
08-NOV-2000; 2000US-0246478P.
08-NOV-2000; 2000US-0246523P.
08-NOV-2000; 2000US-0246524P.
08-NOV-2000; 2000US-0246525P.
08-NOV-2000; 2000US-0246526P.
08-NOV-2000; 2000US-0246527P.
08-NOV-2000; 2000US-0246528P.
08-NOV-2000; 2000US-0246532P.
08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-457726/49.
XX
PT Isolated polypeptide for treating, preventing and prognosing disorders
PT related to the endocrine system including endocrine disorders,
PT reproductive disorders, and gastrointestinal disorders and also for
PT testing and detection e.g. diagnosis.
XX
PS Disclosure; SEQ ID NO 752; 558pp; English.
XX
CC The invention relates to cDNAs encoding novel human endocrine antigens or
CC a fragment having biological activity, a domain, an epitope, full length
CC protein, variant, allelic variant or a species homologue of the
CC cDNA/antigen. The DNAs and polypeptides are useful for preventing,
CC treating or ameliorating a medical condition when administered (e.g. by
CC gene therapy or antisense-therapy). Identifying mutations in the genes
CC coding for the antigens is useful for diagnosing a pathological condition
CC or a susceptibility to a pathological condition. The DNAs, antigens and
CC antibodies raised against the antigens useful for treating, preventing
CC and/ or prognosing disorders related to the endocrine system or hormone
CC imbalance or reproductive disorders, cancers of endocrine tissues,
CC disorders of the pancreas (e.g. diabetes mellitus), the adrenal glands
CC (e.g. hirsutism), ovaries, the thyroid (e.g. hyperthyroidism), the
CC hypothalamus and testes (e.g. vanishing testes syndrome), many examples
CC of diseases and disorders are given in the specification. The present
CC sequence is genomic DNA fragment form a gene encoding an endocrine
CC antigen of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4961 BP; 1245 A; 1244 C; 1408 G; 1064 T; 0 U; 0 Other;

Query Match 17.1%; Score 320.4; DB 4; Length 4961;
Best Local Similarity 82.8%; Pred. No. 2.5e-72;

```
:ches  394;  Conservative    0;  Mismatches   71;  Indels   11;  Gaps    2;

      506 TTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
          ||| |||| ||||| |||||||| ||||  |||||||||||||| ||||| |||||||||
     1770 TTTGTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCTGTAGCCCAAGCTGGAGTG 1711

      566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
          |||||||||||||||| ||| |||||||||  |  |||||||| |||||||  |||||||||
     1710 CAGTGGCATGATCTCGGCCTACTGCAACCACCACCTCCCGGGTTCAAGCAGTTCTCCTGC 1651

      626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATT 685
          |||||||||||||||||||||||||||||||||  | ||||||||||||||| ||||| ||
     1650 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGCACCACCACACCCAGCTATTTTTGTT 1591

      686 TTG---------GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          |||         ||||| |||||||||  ||||| |||||||||||| |||||| |||||
     1590 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 1531

      737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
          ||| | ||||||| |||||| ||||||||||||||||  |||||||||||||| |||| |
     1530 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 1471

      797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
          ||||||| ||  |||||  | |  |||||||     |||||  ||||||||||||||||||
     1470 AGTAGCTGGGCCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 1411

      855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
          |||| |  |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
     1410 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 1351

      915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          |  |||| ||||||||||||||||||||||| ||||| ||||||| |  ||| ||||
     1350 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCCTGCCCGGCC 1295


;T 15
.160_2/c
.nuation (3 of 4) of AEA61160 from base 200001 (Human NEDD4L gene genomic sequence S
:quence split into 4 fragments  LOCUS AEA61160 Accession Aea61160
   Fragment Name      Begin     End
   AEA61160_0             1    110000
   AEA61160_1        100001    210000
   AEA61160_2        200001    310000
   AEA61160_3        300001    354592

:ry Match             17.1%;  Score 320.4;  DB 14;  Length 110000;
:t Local Similarity   80.1%;  Pred. No. 8.6e-72;
:ches  391;  Conservative    0;  Mismatches   91;  Indels    6;  Gaps    1;

      502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
          | ||| |||||  || ||  ||| |||||||||||| |||  |||| | ||  ||| |||
    53549 TGTTTATTTTTATTTATTTATTTATTTTGAGACAGACTCTCACTCTATTGCATAGGGTGG 53490

      562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
          |||||||||||  ||||||| || |||||||||||||  ||||| || ||||||| ||||||
    53489 AGTGCAGTGGCGTGATCTCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCAATTCTC 53430

      622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCC------AGC 675
          ||||||||||||||||||||||||||||||||| |   ||||  |||| |
    53429 GTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGGGAGTGCCAAGACACTCGTTTAACTT 53370

      676 TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGA 735
          ||||| | |||| |||||||||||||||||| |  |||||||||||| |||||||| ||
    53369 TTTTTCTTTTTTTTGAGACAGAGTCTTGCCCTATTGCCCAGGCTGGAGTGCAGTGGCACGA 53310

      736 TCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCT 795
          |||||| |||| || |||||||||||| |||||||||| |||||||||||||| ||||
    53309 TCTCAGCTCACCGCAACCTCCACCTCCCAGGTTCAAGCAGTTCTCCTGCCTCAGCCTCCC 53250

      796 GAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAG 855
          || ||||| | |||| ||  | | ||| |||||  | ||||||||||||||||||||||
    53249 GAATAGCTGGAATTATAGGTGCGTACCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 53190
```

```
Qy        856 AGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGC 915
              |||   | |||  |||||||||||| ||| ||||  ||| || ||||||||||||| |
Db      53189 AGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCCAC 53130

Qy        916 TGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGC 975
              ||||  |||||||||||||| |||||||  |||||||||||||| ||||||  ||||
Db      53129 CTGCCTCAGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCCACAA 53070

Qy        976 TACTCCTT 983
                |||||
Db      53069 GTGTCCTT 53062


Search completed: June 21, 2006, 21:46:16
Job time : 1116 secs
```

GenCore version 5.1.9

ucleic - nucleic search, using sw model

on: June 21, 2006, 21:27:40 ; Search time 10368 Seconds
(without alignments)
11533.727 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ience: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

rched: 6366136 seqs, 31973710525 residues

al number of hits satisfying chosen parameters: 12732272

mum DB seq length: 0
mum DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : GenEmbl:*
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lt o.	Score	% Query Match	Length	DB	ID	Description
1	1854.2	99.2	217470	5	AC040162	AC040162 Homo sapi
2	338.8	18.1	61954	12	AC120915	AC120915 Homo sapi
3	338.8	18.1	63187	5	AL355477	AL355477 Human DNA
4	338.8	18.1	115710	12	AL139015	AL139015 Homo sapi
5	338.8	18.1	149138	12	AC026936	AC026936 Homo sapi
6	335	17.9	128829	5	AC127002	AC127002 Homo sapi
7	333.4	17.8	201508	12	AC026290	AC026290 Homo sapi
8	333	17.8	191353	12	AC090320	AC090320 Homo sapi
9	332.6	17.8	184349	5	AC113189	AC113189 Homo sapi
10	332.6	17.8	187710	5	CNS01DVI	AL135838 Human chr
11	332	17.8	217260	5	AC145375	AC145375 Pan trogl
12	331.2	17.7	88176	12	AL390202_09	Continuation (10 o
13	331.2	17.7	185353	5	AL162732	AL162732 Human DNA
14	331	17.7	101029	5	AC026954	AC026954 Homo sapi
15	331	17.7	121017	5	AC087388	AC087388 Homo sapi
16	331	17.7	166941	12	AC008049	AC008049 Homo sapi
17	331	17.7	222871	12	AC108094	AC108094 Homo sapi
18	330.8	17.7	33362	2	CQ869844	CQ869844 Sequence
c 19	330.8	17.7	92797	5	AF038458	AF038458 Homo sapi
20	330.4	17.7	50303	5	DQ374394	DQ374394 Homo sapi
c 21	330.4	17.7	162991	5	AC027682	AC027682 Homo sapi
22	330.4	17.7	179641	5	AC098591	AC098591 Homo sapi
c 23	330.4	17.7	200618	5	AC009061	AC009061 Homo sapi
c 24	329.8	17.6	39978	5	AC005197	AC005197 Homo sapi
c 25	329.8	17.6	184635	5	AC090239	AC090239 Homo sapi
26	329.8	17.6	186044	12	AC087674	AC087674 Homo sapi
27	329.4	17.6	118684	5	AC025335	AC025335 Homo sapi
28	329	17.6	167906	12	AC164921	AC164921 Pan trogl
c 29	328.8	17.6	188486	12	AP001796	AP001796 Homo sapi
c 30	328.6	17.6	91672	5	AC006468	AC006468 Homo sapi
31	328.6	17.6	218485	12	AC127470	AC127470 Pan trogl
c 32	328.4	17.6	188868	5	AC023232	AC023232 Homo sapi
33	328	17.5	57013	12	AC015769	AC015769 Homo sapi
c 34	327.6	17.5	125756	5	AL590369	AL590369 Human DNA
c 35	327.6	17.5	185511	12	AC007779	AC007779 Homo sapi
c 36	327.4	17.5	136372	5	AC103828	AC103828 Homo sapi
37	327.4	17.5	145450	12	AC068352	AC068352 Homo sapi
38	327.4	17.5	153936	5	AP001207	AP001207 Homo sapi
39	327.4	17.5	166471	12	AC091052	AC091052 Homo sapi
c 40	327	17.5	163542	12	AC129071	AC129071 Pan trogl
41	326.4	17.5	125445	5	AC078795	AC078795 Homo sapi
42	326.4	17.5	131682	5	AL672277	AL672277 Human DNA
c 43	326.4	17.5	148152	12	AC022200	AC022200 Homo sapi
44	326.4	17.5	154563	5	CR936360	CR936360 Human DNA
c 45	326.2	17.4	184857	12	AP002503	AP002503 Homo sapi

ALIGNMENTS

RESULT 1
AC040162
LOCUS AC040162 217470 bp DNA linear PRI 28-MAR-2003
DEFINITION Homo sapiens chromosome 16 clone CTC-479C5, complete sequence.
ACCESSION AC040162
VERSION AC040162.5 GI:29336195
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 217470)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 217470)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 217470)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-APR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 217470)
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Mar 28, 2003 this sequence version replaced gi:19909394.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos National Laboratory
www-shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.

```
URES              Location/Qualifiers
 source           1. .217470
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="16"
                  /clone="CTC-479C5"
IN

ery Match              99.2%;  Score 1854.2;  DB 5;  Length 217470;
st Local Similarity    99.8%;  Pred. No. 0;
tches 1867;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

        1 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCCTGAAATAACTGCTTCT 60
          ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
   166606 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCTTGAAATAACTGCTTCT 166665

       61 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166666 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 166725

      121 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166726 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 166785

      181 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166786 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 166845

      241 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166846 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 166905

      301 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166906 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 166965

      361 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   166966 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 167025

      421 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167026 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 167085

      481 CAGAGATTGAGCTGCTGCTGC-TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGT 539
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
   167086 CAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGT 167145

      540 CTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGC 599
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167146 CTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGC 167205

      600 CTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCA 659
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167206 CTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCA 167265

      660 CGCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTG 719
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167266 CGCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTG 167325

      720 GAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCT 779
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167326 GAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCT 167385

      780 CCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATT 839
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167386 CCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATT 167445

      840 TTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGG 899
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
   167446 TTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGG 167505

Qy        900 ACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCAT 959
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167506 ACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCAC 167565

Qy        960 CGCGCCAGGCCTGAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAAC 1019
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db     167566 CGCGCCAGGCCTGAGCTACTTCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAAC 167625

Qy       1020 GCTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTA 1079
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167626 GCTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTA 167685

Qy       1080 CCATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAG 1139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167686 CCATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAG 167745

Qy       1140 ATAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATA 1199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167746 ATAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATA 167805

Qy       1200 AAGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAA 1259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167806 AAGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAA 167865

Qy       1260 GTGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGG 1319
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167866 GTGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGG 167925

Qy       1320 AGGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGG 1379
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167926 AGGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGG 167985

Qy       1380 TTCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAG 1439
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     167986 TTCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAG 168045

Qy       1440 TCCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTG 1499
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168046 TCCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTG 168105

Qy       1500 CTGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCG 1559
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168106 CTGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCG 168165

Qy       1560 CCTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAAT 1619
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168166 CCTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAAT 168225

Qy       1620 CCCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCC 1679
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168226 CCCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCC 168285

Qy       1680 ACAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACC 1739
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168286 ACAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACC 168345

Qy       1740 TACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCA 1799
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168346 TACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCA 168405

Qy       1800 GGCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGG 1859
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     168406 GGCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGG 168465

Qy       1860 ATCCTTCGTAG 1870
              ||||||||||| 
Db     168466 ATCCTTCGTAG 168476


RESULT 2
AC120915
LOCUS       AC120915               61954 bp    DNA     linear   HTG 23-JUL-2002
DEFINITION  Homo sapiens, *** SEQUENCING IN PROGRESS ***, 22 unordered pieces.
```

```
SSION   AC120915
SION    AC120915.2  GI:21903270
WORDS   HTG; HTGS_PHASE1.
CE      Homo sapiens (human)
GANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
RENCE   1  (bases 1 to 61954)
THORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
        Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
        Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
        Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
        Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
        Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
        Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
        Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
        Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
        Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
        Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
        Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
        Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
        Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
        Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
        Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
        Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
        Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
        Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
        Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
        Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
        Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
        Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
        Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
        Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
        Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
        Moser,M., Neal,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N.,
        Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
        Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
        Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
        Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
        Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
        Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
        Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
        Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S.,
        Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
        Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
        Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
        Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
        Weinstock,G. and Gibbs,R.
TLE     Direct Submission
URNAL   Unpublished
RENCE   2  (bases 1 to 61954)
THORS   Worley,K.C.
TLE     Direct Submission
URNAL   Submitted (14-MAY-2002) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
RENCE   3  (bases 1 to 61954)
THORS   Worley,K.C.
TLE     Direct Submission
URNAL   Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
ENT     On Jul 18, 2002 this sequence version replaced gi:20564251.
        -------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
        -------------- Project Information
            Center project name: GWEB
            Center clone name:
        -------------- Summary Statistics
            Sequencing vector: Plasmid;
            Chemistry: Dye-terminator Big Dye: 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 78665 bases at least Q40
            Consensus quality: 87559 bases at least Q30
            Consensus quality: 93646 bases at least Q20
        --------------
        * NOTE: Estimated insert size may differ from sequence length
        *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 22 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        *        1     2038: contig of 2038 bp in length
        *     2039     2138: gap of unknown length
        *     2139     4239: contig of 2101 bp in length
        *     4240     4339: gap of unknown length
        *     4340     6593: contig of 2254 bp in length
        *     6594     6693: gap of unknown length
        *     6694     8997: contig of 2304 bp in length
        *     8998     9097: gap of unknown length
        *     9098    11920: contig of 2823 bp in length
        *    11921    12020: gap of unknown length
        *    12021    14131: contig of 2111 bp in length
        *    14132    14231: gap of unknown length
        *    14232    16548: contig of 2317 bp in length
        *    16549    16648: gap of unknown length
        *    16649    18948: contig of 2300 bp in length
        *    18949    19048: gap of unknown length
        *    19049    21648: contig of 2600 bp in length
        *    21649    21748: gap of unknown length
        *    21749    24035: contig of 2287 bp in length
        *    24036    24135: gap of unknown length
        *    24136    26191: contig of 2056 bp in length
        *    26192    26291: gap of unknown length
        *    26292    28364: contig of 2073 bp in length
        *    28365    28464: gap of unknown length
        *    28465    31590: contig of 3126 bp in length
        *    31591    31690: gap of unknown length
        *    31691    33847: contig of 2157 bp in length
        *    33848    33947: gap of unknown length
        *    33948    36904: contig of 2957 bp in length
        *    36905    37004: gap of unknown length
        *    37005    40115: contig of 3111 bp in length
        *    40116    40215: gap of unknown length
        *    40216    44221: contig of 4006 bp in length
        *    44222    44321: gap of unknown length
        *    44322    47641: contig of 3320 bp in length
        *    47642    47741: gap of unknown length
        *    47742    50640: contig of 2899 bp in length
        *    50641    50740: gap of unknown length
        *    50741    54553: contig of 3813 bp in length
        *    54554    54653: gap of unknown length
        *    54654    58413: contig of 3760 bp in length
        *    58414    58513: gap of unknown length
        *    58514    61954: contig of 3441 bp in length.
FEATURES             Location/Qualifiers
     source          1. .61954
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gap             2039. .2138
                     /estimated_length=unknown
     gap             4240. .4339
                     /estimated_length=unknown
     gap             6594. .6693
                     /estimated_length=unknown
     gap             8998. .9097
                     /estimated_length=unknown
```

```
gap             11921. .12020
                /estimated_length=unknown
gap             14132. .14231
                /estimated_length=unknown
gap             16549. .16648
                /estimated_length=unknown
gap             18949. .19048
                /estimated_length=unknown
gap             21649. .21748
                /estimated_length=unknown
gap             24036. .24135
                /estimated_length=unknown
gap             26192. .26291
                /estimated_length=unknown
gap             28365. .28464
                /estimated_length=unknown
gap             31591. .31690
                /estimated_length=unknown
gap             33848. .33947
                /estimated_length=unknown
gap             36905. .37004
                /estimated_length=unknown
gap             40116. .40215
                /estimated_length=unknown
gap             44222. .44321
                /estimated_length=unknown
gap             47642. .47741
                /estimated_length=unknown
gap             50641. .50740
                /estimated_length=unknown
gap             54554. .54653
                /estimated_length=unknown
gap             58414. .58513
                /estimated_length=unknown
IN

ery Match             18.1%;  Score 338.8;  DB 12;  Length 61954;
st Local Similarity   83.5%;  Pred. No. 1.5e-96;
tches  396;  Conservative    0;  Mismatches   73;  Indels    5;  Gaps    1;

      503 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 562
          ||  |||||||  |||||  ||||||||||||||||||||  || |||  ||||||||||||
    31983 TTGATTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCACTTTGTTGCCCAGGCTGGA 32042

      563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
          ||  ||||||||||||||  ||  ||||  ||||||||||||||| | ||  |||||||||
    32043 GTACAGTGGCATGATCTGAGCTCACTACAACCTCTGCCTCCCAGGTTGAAGCGATTCAAG 32102

      623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTT 682
          |||||||||||||||||||||||||||||| ||||||||||||||||||    ||| |||||
    32103 TGCCTCAGCCTCCCAAGTAGCTGGGATTTCAGGTGCACGCCACCACACTTGGCTATTTTT 32162

      683 ATTTTG-----GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
           ||||      |||||||||||| || ||||||||  ||||||||||  |||||||   |||
    32163 TTTTTTTTTTTGAGACAGAGTCTAGCTCTGTCACCGAGGCTGGAGTGCAGTGGCGCCATC 32222

      738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
          ||  |||||  |||  |||||  ||||||||||||||||  ||||||||||||||  ||||||||
    32223 TCGGTTCAGTGCAGCCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCATTCTCCTGA 32282

      798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
          ||||||  |||||||||      ||||  |||  |    ||||||||||||||||||||||||
    32283 GTAGCTGGGATTACAGGCACCCACCACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAG 32342

      858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
          |   | ||||||||||||||||  |||  ||||  |||  ||  ||||||||||||||  ||
    32343 ACTGGATTTTGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGATCTGCCT 32402

      918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
          ||||  |||||||||||||||||||| |||  |||||||||  ||  |||  |||||
    32403 GCCTCAGCCTCCCAAAGTGCTGGGATTACAGACGTGAGCCACCGTGCCCGGCCT 32456
```

```
RESULT 3
AL355477
LOCUS       AL355477               63187 bp    DNA     linear   PRI 18-MAY-2005
DEFINITION  Human DNA sequence from clone RP11-113D13 on chromosome 1 Contains
            the 5' end of the MACF1 gene for microtubule-actin crosslinking
            factor 1 and two CpG islands, complete sequence.
ACCESSION   AL355477
VERSION     AL355477.14  GI:16973826
KEYWORDS    HTG; ACTIN; CpG island; MACF1; microtubule.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 63187)
  AUTHORS   Tracey,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vega@sanger.ac.uk
            Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Nov 17, 2001 this sequence version replaced gi:16304908.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            RP11-113D13 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: vega@sanger.ac.uk
            --------------
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one subclone; and the assembly was confirmed by restriction digest,
            except on the rare occasion of the clone being a YAC.
FEATURES             Location/Qualifiers
     source          1. .63187
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /clone="RP11-113D13"
                     /clone_lib="RPCI-11.1"
     misc_feature    1
                     /note="Clone_left_end: RP11-113D13"
     misc_feature    11621. .11840
                     /note="Sequence from AC026936 sequenced by WIBR."
     gene            join(11701. .11928,13748. .13802,14552. .14823,33088. .33366)
                     /gene="MACF1"
                     /locus_tag="RP11-69E11.1-012"
     mRNA            join(11701. .11928,13748. .13802,14552. .14823,33088. .33366)
                     /gene="MACF1"
                     /locus_tag="RP11-69E11.1-012"
                     /product="microtubule-actin crosslinking factor 1"
                     /note="match: ESTs: Em:AA046493.1 Em:AA047436.1
                     Em:BM685084.1 Em:BX107650.1 Em:CD675344.1"
     gene            join(14552. .14823,AL442071.30:5092. .5153,
                     AL442071.30:23924. .24013,AL442071.30:28208. .28285,
                     AL442071.30:31846. .31938,AL442071.30:56118. .56284,
                     AL442071.30:57099. .57211,AL442071.30:57319. .57425,
                     AL442071.30:57976. .58095,AL442071.30:58243. .58338,
```

AL442071.30:58993. .59120,AL442071.30:59207. .59300,
AL442071.30:59514. .59668,AL442071.30:61196. .61472,
AL442071.30:65820. .65948,AL442071.30:66676. .66753,
AL442071.30:67368. .67529,AL442071.30:68956. .69060,
AL442071.30:69697. .69792,AL442071.30:71530. .71655,
AL442071.30:74120. .74272,AL442071.30:76834. .76988,
AL442071.30:83480. .83606,AL442071.30:84155. .84338,
AL442071.30:84741. .84910,AL442071.30:89423. .89601,
AL442071.30:90301. .90439,AL442071.30:91124. .91317,
AL442071.30:92363. .92525,AL442071.30:93574. .93657,
AL442071.30:96518. .96625,AL442071.30:96820. .96918,
AL442071.30:98103. .98255,AL442071.30:98556. .98644,
AL442071.30:101128. .101263,AL137853.12:4093. .4242,
AL137853.12:5400. .5633,AL137853.12:7622. .7846,
AL137853.12:11990. .12508,AL137853.12:13312. .13488,
AL137853.12:13756. .13879,AL137853.12:15408. .15616,
AL137853.12:15975. .16292,AL137853.12:22763. .22929,
AL137853.12:24645. .24801,AL137853.12:27094. .27211,
AL137853.12:33076. .33138,AL137853.12:33818. .34059,
AL137853.12:34859. .34988,AL137853.12:36568. .36767,
AL137853.12:40049. .40544,AL137853.12:41802. .43273,
AL137853.12:76989. .77150,AL137853.12:77345. .77515,
AL137853.12:78643. .78809,AL137853.12:82070. .82229,
AL137853.12:82576. .82902,AL137853.12:83857. .83941,
AL137853.12:84442. .84683,AL137853.12:85317. .85523,
AL137853.12:87150. .87269,AL137853.12:87663. .87770,
AL137853.12:89029. .89247,AL137853.12:90173. .90376,
AL137853.12:91097. .91105,AL137853.12:92392. .92623,
AL137853.12:93889. .94115,AL137853.12:95618. .95740,
AL137853.12:95839. .95940,AL137853.12:96567. .96668,
AL137853.12:96762. .96935,AL137853.12:97027. .97182,
AL137853.12:97329. .97526,AL137853.12:98061. .98189,
AL137853.12:99277. .99450,AL137853.12:102047. .102202,
AL137853.12:102307. .102513,AL137853.12:102648. .102767,
AL137853.12:103209. .103378,AL137853.12:105667. .105826,
AL137853.12:106771. .106965,AL137853.12:107299. .107427,
AL137853.12:108322. .108465,AL137853.12:109524. .109712,
AL137853.12:112985. .113105,AL137853.12:113701. .113853,
AL365277.23:2001. .2011,AL365277.23:3394. .3556,
AL365277.23:4645. .4799,AL365277.23:6374. .6448,
AL365277.23:11377. .11494,AL365277.23:22560. .22771,
AL365277.23:27363. .27492,AL365277.23:28300. .29749)
/gene="MACF1"
/locus_tag="RP11-69E11.1-001"

mRNA join(14552. .14823,AL442071.30:5092. .5153,
AL442071.30:23924. .24013,AL442071.30:28208. .28285,
AL442071.30:31846. .31938,AL442071.30:56118. .56284,
AL442071.30:57099. .57211,AL442071.30:57319. .57425,
AL442071.30:57976. .58095,AL442071.30:58243. .58338,
AL442071.30:58993. .59120,AL442071.30:59207. .59300,
AL442071.30:59514. .59668,AL442071.30:61196. .61472,
AL442071.30:65820. .65948,AL442071.30:66676. .66753,
AL442071.30:67368. .67529,AL442071.30:68956. .69060,
AL442071.30:69697. .69792,AL442071.30:71530. .71655,
AL442071.30:74120. .74272,AL442071.30:76834. .76988,
AL442071.30:83480. .83606,AL442071.30:84155. .84338,
AL442071.30:84741. .84910,AL442071.30:89423. .89601,
AL442071.30:90301. .90439,AL442071.30:91124. .91317,
AL442071.30:92363. .92525,AL442071.30:93574. .93657,
AL442071.30:96518. .96625,AL442071.30:96820. .96918,
AL442071.30:98103. .98255,AL442071.30:98556. .98644,
AL442071.30:101128. .101263,AL137853.12:4093. .4242,
AL137853.12:5400. .5633,AL137853.12:7622. .7846,
AL137853.12:11990. .12508,AL137853.12:13312. .13488,
AL137853.12:13756. .13879,AL137853.12:15408. .15616,
AL137853.12:15975. .16292,AL137853.12:22763. .22929,
AL137853.12:24645. .24801,AL137853.12:27094. .27211,
AL137853.12:33076. .33138,AL137853.12:33818. .34059,
AL137853.12:34859. .34988,AL137853.12:36568. .36767,
AL137853.12:40049. .40544,AL137853.12:41802. .43273,
AL137853.12:76989. .77150,AL137853.12:77345. .77515,
AL137853.12:78643. .78809,AL137853.12:82070. .82229,
AL137853.12:82576. .82902,AL137853.12:83857. .83941,
AL137853.12:84442. .84683,AL137853.12:85317. .85523,
AL137853.12:87150. .87269,AL137853.12:87663. .87770,
AL137853.12:89029. .89247,AL137853.12:90173. .90376,
AL137853.12:91097. .91105,AL137853.12:92392. .92623,
AL137853.12:93889. .94115,AL137853.12:95618. .95740,
AL137853.12:95839. .95940,AL137853.12:96567. .96668,
AL137853.12:96762. .96935,AL137853.12:97027. .97182,
AL137853.12:97329. .97526,AL137853.12:98061. .98189,
AL137853.12:99277. .99450,AL137853.12:102047. .102202,
AL137853.12:102307. .102513,AL137853.12:102648. .102767,
AL137853.12:103209. .103378,AL137853.12:105667. .105826,
AL137853.12:106771. .106965,AL137853.12:107299. .107427,
AL137853.12:108322. .108465,AL137853.12:109524. .109712,
AL137853.12:112985. .113105,AL137853.12:113701. .113853,
AL365277.23:2001. .2011,AL365277.23:3394. .3556,
AL365277.23:4645. .4799,AL365277.23:6374. .6448,
AL365277.23:11377. .11494,AL365277.23:22560. .22771,
AL365277.23:27363. .27492,AL365277.23:28300. .29749)
/gene="MACF1"
/locus_tag="RP11-69E11.1-001"
/product="microtubule-actin crosslinking factor 1"
/note="match: ESTs: Em:AA169873.1 Em:AF150358.1
Em:AL138426.1 Em:AL698142.1 Em:AL699066.1 Em:AL710538.1
Em:AL713492.1 Em:AU134479.1 Em:AW043677.1 Em:AY032902.1
Em:BE007711.1 Em:BE396321.1 Em:BE767805.1 Em:BE771881.1
Em:BE837593.1 Em:BE838157.1 Em:BE843469.1 Em:BF025979.1
Em:BG751834.1 Em:BG993158.1 Em:BI600856.1 Em:BI769261.1
Em:BM471359.1 Em:BQ014846.1 Em:BQ058523.1 Em:BQ128354.1
Em:BQ309642.1 Em:BQ328711.1 Em:BU625288.1 Em:BX490007.1
Em:BX493047.1 Em:BX508137.1 Em:BX646635.1 Em:CA390757.1
Em:CB045418.1 Em:CB051751.1 Em:CB051752.1 Em:CB267890.1
Em:CB850509.1 Em:CD106951.1 Em:CD643646.1 Em:CD676309.1
Em:H13461.1 Em:W61052.1 Em:Z21839.1
match: cDNAs: Em:AB029290.1 Em:AK122291.1 Em:BC026707.1"

CDS join(14604. .14823,AL442071.30:5092. .5153,
AL442071.30:23924. .24013,AL442071.30:28208. .28285,
AL442071.30:31846. .31938,AL442071.30:56118. .56284,
AL442071.30:57099. .57211,AL442071.30:57319. .57425,
AL442071.30:57976. .58095,AL442071.30:58243. .58338,
AL442071.30:58993. .59120,AL442071.30:59207. .59300,
AL442071.30:59514. .59668,AL442071.30:61196. .61472,
AL442071.30:65820. .65948,AL442071.30:66676. .66753,
AL442071.30:67368. .67529,AL442071.30:68956. .69060,
AL442071.30:69697. .69792,AL442071.30:71530. .71655,
AL442071.30:74120. .74272,AL442071.30:76834. .76988,
AL442071.30:83480. .83606,AL442071.30:84155. .84338,
AL442071.30:84741. .84910,AL442071.30:89423. .89601,
AL442071.30:90301. .90439,AL442071.30:91124. .91317,
AL442071.30:92363. .92525,AL442071.30:93574. .93657,
AL442071.30:96518. .96625,AL442071.30:96820. .96918,
AL442071.30:98103. .98255,AL442071.30:98556. .98644,
AL442071.30:101128. .101263,AL137853.12:4093. .4242,
AL137853.12:5400. .5633,AL137853.12:7622. .7846,
AL137853.12:11990. .12508,AL137853.12:13312. .13488,
AL137853.12:13756. .13879,AL137853.12:15408. .15616,
AL137853.12:15975. .16292,AL137853.12:22763. .22929,

Query Match 18.1%; Score 338.8; DB 5; Length 63187;
Best Local Similarity 80.8%; Pred. No. 1.6e-96;
Matches 409; Conservative 0; Mismatches 92; Indels 5; Gaps 1;

```
Qy       507 TTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGC 566
             ||| ||| ||| | |  |||||||||||  |||||| ||| ||| ||||||||||||||||
Db      8384 TTTATTTGTTTGTTTCCTTTTTTGAGATGGAGTCTCGCTTTGTTGCCCAGGCTGGAGTGC 8443

Qy       567 AGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCC 626
             ||||||| | |||| || |||||||||||| ||||||| | ||||||||||||||||||||
Db      8444 AGTGGCACGGTCTCAGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCGATTCTCCTGCG 8503

Qy       627 TCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTT 681
             |||||||||  |||||||||||||||||  || ||||||||||||||     ||||||
Db      8504 TCAGCCTCCTGAGTAGCTGGGATTACAGGCACATGCCACCACACCCAGCTAATTTTTTTT 8563
```

```
 682 TATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAG 741
     | |||| ||||  |||||||| ||||| |||||||||||| |||||||||||||||||
8564 TTTTTTTGAGATGGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCAG 8623

 742 TTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAG 801
     ||||||| |||||| |||||||||||||||| |||||||||||||| || ||||||||
8624 CTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTACTGAGTAG 8683

 802 CTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGC 861
     || |||||||||  |||  || ||||   |||||||||||||||||||||||||||| |
8684 CTGGGATTACAGGTGTGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGG 8743

 862 GCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCT 921
     | |||  |||||||| ||| ||| |||| ||||||| || |||||||||| ||    ||
8744 GGTTTCACCATGTTGACCAGGCTGGTCTCGAACCCCTGAACTCAGGTGATCTGCCCATCT 8803

 922 TGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCC 981
       ||| ||||||||||||||||| |||| |||| || ||| || ||||   ||   ||
8804 CTGCCACCCAAAGTGCTGGGATTACAGGTCTGAGACACCGCACCCGGCCTACTCTCTGCC 8863

 982 TTTAGTCTCTGGAAAGACTGCGGCTC 1007
     | |||    ||  | |   |  || ||
8864 TCTATGAGCTCAATATTTTTAGGTTC 8889
```

```
LT 4
9015/c
S       AL139015              115710 bp    DNA     linear   HTG 06-SEP-2001
NITION  Homo sapiens chromosome 1 clone RP4-648J17 map p34.1-34.3, 6
        unordered pieces.
SSION   AL139015
ION     AL139015.6  GI:13567858
ORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
CE      Homo sapiens (human)
GANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
RENCE   1
THORS   Plumb,B.
TLE     Direct Submission
URNAL   Submitted (05-SEP-2001) Sanger Centre, Hinxton, Cambridgeshire,
        CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
        requests: clonerequest@sanger.ac.uk
ENT     On Apr 9, 2001 this sequence version replaced gi:9796296.
        -------------- Genome Center
        Center: Sanger Centre
        Center code: SC
        Web site: http://www.sanger.ac.uk
        Contact: humquery@sanger.ac.uk
        -------------- Project Information
        Center project name: dJ648J17
        -------------- Summary Statistics
        Assembly program: XGAP4; version 4.5
        Sequencing vector: plasmid; L08752; 100% of reads
        Chemistry: Dye-terminator Big Dye; 100% of reads
        Consensus quality: 113662 bases at least Q40
        Consensus quality: 114216 bases at least Q30
        Consensus quality: 114601 bases at least Q20
        Insert size: 115210; sum-of-contigs
        Insert size: 132500; 15.0% error; agarose-fp
        Quality coverage: 6.13x in Q20 bases; sum-of-contigs Quality
        coverage: 5.53x in Q20 bases; agarose-fp
        --------------.
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 6 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        *       1    36330: contig of 36330 bp in length
        *    36331    36430: gap of 100 bp
        *    36431    47455: contig of 11025 bp in length
        *    47456    47555: gap of 100 bp
        *    47556    52799: contig of 5244 bp in length
        *    52800    52899: gap of 100 bp
        *    52900    60268: contig of 7369 bp in length
        *    60269    60368: gap of 100 bp
        *    60369    88988: contig of 28620 bp in length
        *    88989    89088: gap of 100 bp
        *    89089   115710: contig of 26622 bp in length.
FEATURES             Location/Qualifiers
     source          1. .115710
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p34.1-34.3"
                     /clone="RP4-648J17"
                     /clone_lib="RPCI-4"
     misc_feature    1. .36330
                     /note="assembly_fragment:01504
                     fragment_chain:1"
     misc_feature    36431. .47455
                     /note="assembly_fragment:01208
                     fragment_chain:1"
     misc_feature    47556. .52799
                     /note="assembly_fragment:00376
                     fragment_chain:1"
     misc_feature    52900. .60268
                     /note="assembly_fragment:01176
                     fragment_chain:1"
     misc_feature    60369. .88988
                     /note="assembly_fragment:00586
                     fragment_chain:1"
     misc_feature    89089. .115710
                     /note="assembly_fragment:00793.0"
ORIGIN
```

```
  Query Match             18.1%;  Score 338.8;  DB 12;  Length 115710;
  Best Local Similarity   80.8%;  Pred. No. 1.8e-96;
  Matches  409;  Conservative    0;  Mismatches   92;  Indels    5;  Gaps    1;

Qy       507 TTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGC 566
             ||| ||| | |  |||||||||  |||||| ||| ||| ||||||||||||||||||
Db      3046 TTTATTTGTTTGTTTCCTTTTTTGAGATGGAGTCTCGCTTTGTAGCCCAGGCTGGAGTGC 2987

Qy       567 AGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCC 626
             ||||||| | |||| || |||||||||||| ||||||| | ||||||||||||||||||
Db      2986 AGTGGCACGGTCTCAGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCGATTCTCCTGCG 2927

Qy       627 TCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTT 681
             |||||||||  |||||||||||||||||||  || ||||||||||||||     ||||||
Db      2926 TCAGCCTCCTGAGTAGCTGGGATTACAGGCACATGCCACCACACCCAGCTAATTTTTTTT 2867

Qy       682 TATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAG 741
             | |||| ||||  |||||||| ||||| |||||||||||| |||||||||||||||||
Db      2866 TTTTTTTGAGATGGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCAG 2807

Qy       742 TTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAG 801
             ||||||| |||||| |||||||||||||||| |||||||||||||| || ||||||||
Db      2806 CTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTACTGAGTAG 2747

Qy       802 CTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGC 861
             || |||||||||  |||  || ||||   |||||||||||||||||||||||||||| |
Db      2746 CTGGGATTACAGGTGTGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGG 2687

Qy       862 GCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCT 921
             | |||  |||||||| ||| ||| |||| ||||||| || |||||||||| ||    ||
Db      2686 GGTTTCACCATGTTGACCAGGCTGGTCTCGAACCCCTGAACTCAGGTGATCTGCCCATCT 2627

Qy       922 TGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCC 981
               ||| ||||||||||||||||| |||| |||| || ||| || |||||   ||   ||
Db      2626 CTGCCACCCAAAGTGCTGGGATTACAGGTCTGAGACACCGCACCCGGCCTACTCTCTGCC 2567
```

```
      982 TTTAGTCTCTGGAAAGACTGCGGCTC 1007
          | ||    ||  | |   |  || ||
     2566 TCTATGAGCTCAATATTTTAGGTTC 2541


:LT 5
:6936
)S        AC026936              149138 bp    DNA     linear   HTG 27-APR-2000
:NITION   Homo sapiens chromosome 1 clone RP11-186C2 map 1, WORKING DRAFT
          SEQUENCE, 33 unordered pieces.
:SSION    AC026936
:ION      AC026936.2  GI:7652013
'ORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
:CE       Homo sapiens (human)
:GANISM   Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Homo.
:RENCE    1  (bases 1 to 149138)
'THORS    Birren,B., Linton,L., Nusbaum,C. and Lander,E.
'TLE      Homo sapiens chromosome 1, clone RP11-186C2
'URNAL    Unpublished
:RENCE    2  (bases 1 to 149138)
:THORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
          Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
          Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
          Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
          Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
          Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
          Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zimmer,A. and Zody,M.
:TLE      Direct Submission
)URNAL    Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
'ENT      On Apr 27, 2000 this sequence version replaced gi:7328801.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
          -------------- Project Information
              Center project name: L7617
              Center clone name: 186_C_2
          -------------- Summary Statistics
              Sequencing vector: M13; M77815; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 132550 bases at least Q40
              Consensus quality: 140551 bases at least Q30
              Consensus quality: 144026 bases at least Q20
              Insert size: 157000; agarose-fp
              Insert size: 145938; sum-of-contigs
              Quality coverage: 4.0 in Q20 bases; agarose-fp
              Quality coverage: 4.3 in Q20 bases; sum-of-contigs
          --------------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 33 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1     1364: contig of 1364 bp in length
          *     1365     1464: gap of 100 bp
          *     1465     2866: contig of 1402 bp in length
          *     2867     2966: gap of 100 bp
          *     2967     4325: contig of 1359 bp in length
          *     4326     4425: gap of 100 bp
          *     4426     5455: contig of 1030 bp in length
          *     5456     5555: gap of 100 bp
          *     5556     6761: contig of 1206 bp in length
          *     6762     6861: gap of 100 bp
          *     6862     8299: contig of 1438 bp in length
          *     8300     8399: gap of 100 bp
          *     8400    11139: contig of 2740 bp in length
          *    11140    11239: gap of 100 bp
          *    11240    13012: contig of 1773 bp in length
          *    13013    13112: gap of 100 bp
          *    13113    15320: contig of 2208 bp in length
          *    15321    15420: gap of 100 bp
          *    15421    16828: contig of 1408 bp in length
          *    16829    16928: gap of 100 bp
          *    16929    18769: contig of 1841 bp in length
          *    18770    18869: gap of 100 bp
          *    18870    20670: contig of 1801 bp in length
          *    20671    20770: gap of 100 bp
          *    20771    22998: contig of 2228 bp in length
          *    22999    23098: gap of 100 bp
          *    23099    23754: contig of 656 bp in length
          *    23755    23854: gap of 100 bp
          *    23855    26081: contig of 2227 bp in length
          *    26082    26181: gap of 100 bp
          *    26182    28819: contig of 2638 bp in length
          *    28820    28919: gap of 100 bp
          *    28920    30950: contig of 2031 bp in length
          *    30951    31050: gap of 100 bp
          *    31051    34127: contig of 3077 bp in length
          *    34128    34227: gap of 100 bp
          *    34228    37535: contig of 3308 bp in length
          *    37536    37635: gap of 100 bp
          *    37636    40364: contig of 2729 bp in length
          *    40365    40464: gap of 100 bp
          *    40465    45668: contig of 5204 bp in length
          *    45669    45768: gap of 100 bp
          *    45769    51359: contig of 5591 bp in length
          *    51360    51459: gap of 100 bp
          *    51460    56834: contig of 5375 bp in length
          *    56835    56934: gap of 100 bp
          *    56935    61504: contig of 4570 bp in length
          *    61505    61604: gap of 100 bp
          *    61605    67102: contig of 5498 bp in length
          *    67103    67202: gap of 100 bp
          *    67203    72252: contig of 5050 bp in length
          *    72253    72352: gap of 100 bp
          *    72353    78175: contig of 5823 bp in length
          *    78176    78275: gap of 100 bp
          *    78276    83342: contig of 5067 bp in length
          *    83343    83442: gap of 100 bp
          *    83443    91033: contig of 7591 bp in length
          *    91034    91133: gap of 100 bp
          *    91134    98803: contig of 7670 bp in length
          *    98804    98903: gap of 100 bp
          *    98904   113936: contig of 15033 bp in length
          *   113937   114036: gap of 100 bp
          *   114037   129279: contig of 15243 bp in length
          *   129280   129379: gap of 100 bp
          *   129380   149138: contig of 19759 bp in length.
FEATURES             Location/Qualifiers
```

```
source          1. .149138
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="1"
                /map="1"
                /clone="RP11-186C2"
                /clone_lib="RPCI-11 Human Male BAC"
misc_feature    1. .1364
                /note="assembly_fragment"
gap             1365. .1464
                /estimated_length=100
misc_feature    1465. .2866
                /note="assembly_fragment"
gap             2867. .2966
                /estimated_length=100
misc_feature    2967. .4325
                /note="assembly_fragment"
gap             4326. .4425
                /estimated_length=100
misc_feature    4426. .5455
                /note="assembly_fragment"
gap             5456. .5555
                /estimated_length=100
misc_feature    5556. .6761
                /note="assembly_fragment"
gap             6762. .6861
                /estimated_length=100
misc_feature    6862. .8299
                /note="assembly_fragment"
gap             8300. .8399
                /estimated_length=100
misc_feature    8400. .11139
                /note="assembly_fragment"
gap             11140. .11239
                /estimated_length=100
misc_feature    11240. .13012
                /note="assembly_fragment"
gap             13013. .13112
                /estimated_length=100
misc_feature    13113. .15320
                /note="assembly_fragment"
gap             15321. .15420
                /estimated_length=100
misc_feature    15421. .16828
                /note="assembly_fragment"
gap             16829. .16928
                /estimated_length=100
misc_feature    16929. .18769
                /note="assembly_fragment"
gap             18770. .18869
                /estimated_length=100
misc_feature    18870. .20670
                /note="assembly_fragment"
gap             20671. .20770
                /estimated_length=100
misc_feature    20771. .22998
                /note="assembly_fragment"
gap             22999. .23098
                /estimated_length=100
misc_feature    23099. .23754
                /note="assembly_fragment

ery Match             18.1%;  Score 338.8;  DB 12;  Length 149138;
st Local Similarity   80.8%;  Pred. No. 1.9e-96;
tches  409;  Conservative    0;  Mismatches   92;  Indels    5;  Gaps    1;

      507 TTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGC 566
          ||| ||| ||| | |  |||||||||  |||||| ||| ||| ||||||||||||||||
   133048 TTTATTTGTTTGTTTCCTTTTTTGAGATGGAGTCTCGCTTTGTTGCCCAGGCTGGAGTGC 133107

      567 AGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCC 626
          ||||||| | |||| || |||||||||||| ||||||| | ||||||||||||||||||
```

```
Db     133108 AGTGGCACGGTCTCAGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCGATTCTCCTGCG 133167

Qy        627 TCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTT 681
              |||||||||  |||||||||||||||||  || ||||||||||||||     ||||||
Db     133168 TCAGCCTCCTGAGTAGCTGGGATTACAGGCACATGCCACCACACCCAGCTAATTTTTTTT 133227

Qy        682 TATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAG 741
              | |||| ||||  |||||||| |||| |||||||||||| ||||||||||||||||||
Db     133228 TTTTTTTGAGATGGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCAG 133287

Qy        742 TTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAG 801
              |||||| |||||| |||||||||||||||| |||||||||||||| || ||||||||
Db     133288 CTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTACTGAGTAG 133347

Qy        802 CTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGC 861
              || |||||||||  |||  || ||||   |||||||||||||||||||||||||||| |
Db     133348 CTGGGATTACAGGTGTGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGG 133407

Qy        862 GCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCT 921
              | |||  |||||||| ||| ||| |||| ||||||| || ||||||||||| ||    ||
Db     133408 GGTTTCACCATGTTGACCAGGCTGGTCTCGAACCCCTGAACTCAGGTGATCTGCCCATCT 133467

Qy        922 TGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCC 981
              |||  |||||||||||||||| ||||  |||| || ||| || |||||   ||   ||
Db     133468 CTGCCACCCAAAGTGCTGGGATTACAGGTCTGAGACACCGCACCCGGCCTACTCTCTGCC 133527

Qy        982 TTTAGTCTCTGGAAAGACTGCGGCTC 1007
              | ||    ||  | |   |  || ||
Db     133528 TCTATGAGCTCAATATTTTAGGTTC 133553
```

```
RESULT 6
AC127002
LOCUS       AC127002              128829 bp    DNA     linear   PRI 15-MAR-2003
DEFINITION  Homo sapiens 12 BAC RP11-15J22 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC127002
VERSION     AC127002.3  GI:25281341
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 128829)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
            Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
            Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,
            Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
            Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P.,
            Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M.,
            Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S.,
            Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
            Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
            King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
            Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,
            Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X.,
            Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P.,
            Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
            Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,
            Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
            Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
```

Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.

:TLE Direct Submission
)URNAL Unpublished
:RENCE 2 (bases 1 to 128829)
/THORS Worley,K.C.
:TLE Direct Submission
)URNAL Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
:RENCE 3 (bases 1 to 128829)
/THORS Worley,K.C.
:TLE Direct Submission
)URNAL Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
:RENCE 4 (bases 1 to 128829)
/THORS Worley,K.C.
:TLE Direct Submission
:URNAL Submitted (25-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
:RENCE 5 (bases 1 to 128829)
/THORS Worley,K.C.
:TLE Direct Submission
'URNAL Submitted (23-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
:RENCE 6 (bases 1 to 128829)
/THORS Worley,K.C.
'TLE Direct Submission
:URNAL Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
'ENT On Nov 25, 2002 this sequence version replaced gi:23343647.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES	Location/Qualifiers
source	1. .128829
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-15J22"
misc_feature	complement(1. .2000)
	/note="overlaps bases 1. .2000 of clone AC093016"
	/function="clone overlap"
repeat_region	4. .248
	/rpt_family="AluSg/x"
repeat_region	499. .791
	/rpt_family="AluSx"
repeat_region	905. .1203
	/rpt_family="AluSq"
repeat_region	1499. .1779
	/rpt_family="AluSx"
repeat_region	1780. .1809
	/rpt_family="(ATTG)n"
repeat_region	1877. .2102
	/rpt_family="MIR"
repeat_region	3349. .3436
	/rpt_family="(CCG)n"
repeat_region	3477. .3513
	/rpt_family="GC_rich"
repeat_region	3529. .3643
	/rpt_family="(CCCGG)n"
repeat_region	4895. .5183
	/rpt_family="AluSx"
repeat_region	5184. .5248
	/rpt_family="(TA)n"
repeat_region	complement(5350. .5655)
	/rpt_family="AluY"
repeat_region	5656. .5683
	/rpt_family="AT_rich"
repeat_region	5696. .5945
	/rpt_family="AluSg"
misc_feature	5923. .5925
	/function="low quality"
repeat_region	complement(6120. .6421)
	/rpt_family="AluY"
repeat_region	6703. .6755
	/rpt_family="HAL1"
repeat_region	complement(6756. .7068)
	/rpt_family="AluJb"
repeat_region	7069. .7196
	/rpt_family="HAL1"
repeat_region	complement(7237. .7544)
	/rpt_family="AluJo"
repeat_region	7598. .7738
	/rpt_family="AluJo"
repeat_region	7739. .8015
	/rpt_family="AluSx"
repeat_region	8016. .8056
	/rpt_family="(TAAA)n"
repeat_region	8057. .8225
	/rpt_family="AluJo"
repeat_region	8242. .8264
	/rpt_family="(TTG)n"
repeat_region	complement(8273. .8385)
	/rpt_family="FLAM_C"

```
repeat_region   complement(8389. .8698)
                /rpt_family="AluSx"
repeat_region   complement(8927. .9215)
                /rpt_family="AluSg"
repeat_region   9587. .9612
                /rpt_family="(TTTTG)n"
repeat_region   complement(9613. .9898)
                /rpt_family="AluSq"
repeat_region   9918. .10185
                /rpt_family="AluSx"
repeat_region   10338. .10478
                /rpt_family="AluJb"
repeat_region   10559. .10657
                /rpt_family="L1MA8"
```

```
ery Match             17.9%;  Score 335;  DB 5;  Length 128829;
st Local Similarity   81.8%;  Pred. No. 3.1e-95;
tches  400;  Conservative    0;  Mismatches   85;  Indels    4;  Gaps    1;

      502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
          ||| ||| |||  || ||  ||| ||||||||||||||||  |||||||  ||||||||||
    28862 TTTATTTATTTATTTATTTATTTTTGAGACAGAGTCTCACTCTGTCTCCCAGGCTGG 28921

      562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
          ||| ||||||||||||||  || ||||||||||||||||||| | | ||||||  |||||||
    28922 AGTTCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCTCAGGTTCAAGTGATTCTC 28981

      622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT----T 677
          |||||||||||||  ||||||||  |||||||| |  | ||||||||||  |||    |
    28982 CTGCCTCAGCCTCCTGAGTAGCTGCAATTACAGGCGTGCACCACCACACCTGGCTAATTT 29041

      678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
          ||||| |||| ||||  |||||| || ||||  |||||| |||||| |||||| |||||
    29042 TTTTTTTTTTTGAGATGGAGTCTCGCTCTGTTGCCCAGGTTGGAGTGCAGTGGCGTGATC 29101

      738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
          || | |||||||  ||||||||||| ||||||||| |||||||||||||| ||||  |
    29102 TCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCTA 29161

      798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
          |||||| ||| |||||  |||| || ||| |  |||||||||||||||||||||||||
    29162 GTAGCTGGGACTACAGGCGTGCGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAG 29221

      858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
          ||| | |||  |||||||||||| ||| |||| |||| || |||||| |||||||||
    29222 ATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCC 29281

      918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTA 977
           |||  |||||||||||||||||||| ||| ||||||||| || ||| |||||    |
    29282 ACCTCAGCCTCCCAAAGTGCTGGGATTTCAGCCGTGAGCCACCGTGCCCGGCCTATTTTT 29341

      978 CTCCTTTAG 986
          |  |||||
    29342 ATTTTTTAG 29350
```

```
LT 7
6290
S       AC026290              201508 bp    DNA     linear   HTG 07-JUL-2000
NITION  Homo sapiens chromosome 12 clone RP11-657I11, WORKING DRAFT
        SEQUENCE, 44 unordered pieces.
SION    AC026290
ON      AC026290.2  GI:8516761
RDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
E       Homo sapiens (human)
ANISM   Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
RENCE   1  (bases 1 to 201508)
THORS   Waterston,R.H.
TLE     The sequence of Homo sapiens clone
URNAL   Unpublished
```

```
REFERENCE   2  (bases 1 to 201508)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 14, 2000 this sequence version replaced gi:7272316.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0657I11
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 183180 bases at least Q40
            Consensus quality: 188526 bases at least Q30
            Consensus quality: 190843 bases at least Q20
            Insert size: 223000; agarose-fp
            Insert size: 197208; sum-of-contigs
            Quality coverage: 3.45 in Q20 bases; agarose-fp
            Quality coverage: 3.87 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 44 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1334: contig of 1334 bp in length
            *     1335     1434: gap of unknown length
            *     1435     2470: contig of 1036 bp in length
            *     2471     2570: gap of unknown length
            *     2571     3685: contig of 1115 bp in length
            *     3686     3785: gap of unknown length
            *     3786     5164: contig of 1379 bp in length
            *     5165     5264: gap of unknown length
            *     5265     7133: contig of 1869 bp in length
            *     7134     7233: gap of unknown length
            *     7234     8409: contig of 1176 bp in length
            *     8410     8509: gap of unknown length
            *     8510    10159: contig of 1650 bp in length
            *    10160    10259: gap of unknown length
            *    10260    11978: contig of 1719 bp in length
            *    11979    12078: gap of unknown length
            *    12079    13423: contig of 1345 bp in length
            *    13424    13523: gap of unknown length
            *    13524    14829: contig of 1306 bp in length
            *    14830    14929: gap of unknown length
            *    14930    17244: contig of 2315 bp in length
            *    17245    17344: gap of unknown length
            *    17345    19950: contig of 2606 bp in length
            *    19951    20050: gap of unknown length
            *    20051    23459: contig of 3409 bp in length
            *    23460    23559: gap of unknown length
            *    23560    26491: contig of 2932 bp in length
            *    26492    26591: gap of unknown length
            *    26592    29687: contig of 3096 bp in length
            *    29688    29787: gap of unknown length
            *    29788    32790: contig of 3003 bp in length
            *    32791    32890: gap of unknown length
            *    32891    35373: contig of 2483 bp in length
            *    35374    35473: gap of unknown length
            *    35474    38408: contig of 2935 bp in length
            *    38409    38508: gap of unknown length
            *    38509    41522: contig of 3014 bp in length
```

```
        *    41523    41622: gap of unknown length
        *    41623    44227: contig of 2605 bp in length
        *    44228    44327: gap of unknown length
        *    44328    46523: contig of 2196 bp in length
        *    46524    46623: gap of unknown length
        *    46624    49080: contig of 2457 bp in length
        *    49081    49180: gap of unknown length
        *    49181    51688: contig of 2508 bp in length
        *    51689    51788: gap of unknown length
        *    51789    54581: contig of 2793 bp in length
        *    54582    54681: gap of unknown length
        *    54682    58182: contig of 3501 bp in length
        *    58183    58282: gap of unknown length
        *    58283    62871: contig of 4589 bp in length
        *    62872    62971: gap of unknown length
        *    62972    66745: contig of 3774 bp in length
        *    66746    66845: gap of unknown length
        *    66846    71014: contig of 4169 bp in length
        *    71015    71114: gap of unknown length
        *    71115    74743: contig of 3629 bp in length
        *    74744    74843: gap of unknown length
        *    74844    79286: contig of 4443 bp in length
        *    79287    79386: gap of unknown length
        *    79387    83407: contig of 4021 bp in length
        *    83408    83507: gap of unknown length
        *    83508    88267: contig of 4760 bp in length
        *    88268    88367: gap of unknown length
        *    88368    93902: contig of 5535 bp in length
        *    93903    94002: gap of unknown length
        *    94003    99104: contig of 5102 bp in length
        *    99105    99204: gap of unknown length
        *    99205   103797: contig of 4593 bp in length
        *   103798   103897: gap of unknown length
        *   103898   110710: contig of 6813 bp in length
        *   110711   110810: gap of unknown length
        *   110811   117968: contig of 7158 bp in length
        *   117969   118068: gap of unknown length
        *   118069   123971: contig of 5903 bp in length
        *   123972   124071: gap of unknown length
        *   124072   132237: contig of 8166 bp in length
        *   132238   132337: gap of unknown length
        *   132338   140133: contig of 7796 bp in length
        *   140134   140233: gap of unknown length
        *   140234   151490: contig of 11257 bp in length
        *   151491   151590: gap of unknown length
        *   151591   164136: contig of 12546 bp in length
        *   164137   164236: gap of unknown length
        *   164237   174578: contig of 10342 bp in length
        *   174579   174678: gap of unknown length
        *   174679   201508: contig of 26830 bp in length.
FEATURES             Location/Qualifiers
     source          1. .201508
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="RP11-657I11"
     misc_feature    1. .1334
                     /note="assembly_name:Contig11"
     gap             1335. .1434
                     /estimated_length=unknown
     misc_feature    1435. .2470
                     /note="assembly_name:Contig12"
     gap             2471. .2570
                     /estimated_length=unknown
     misc_feature    2571. .3685
                     /note="assembly_name:Contig13"
     gap             3686. .3785
                     /estimated_length=unknown
     misc_feature    3786. .5164
                     /note="assembly_name:Contig14"
     gap             5165. .5264
                     /estimated_length=unknown
     misc_feature    5265. .7133
                     /note="assembly_name:Contig15"
     gap             7134. .7233
                     /estimated_length=unknown
     misc_feature    7234. .8409
                     /note="assembly_name:Contig16"
     gap             8410. .8509
                     /estimated_length=unknown
     misc_feature    8510. .10159
                     /note="assembly_name:Contig17"
     gap             10160. .10259
                     /estimated_length=unknown
     misc_feature    10260. .11978
                     /note="assembly_name:Contig18"
     gap             11979. .12078
                     /estimated_length=unknown
     misc_feature    12079. .13423
                     /note="assembly_name:Contig19"
     gap             13424. .13523
                     /estimated_length=unknown
     misc_feature    13524. .14829
                     /note="assembly_name:Contig20"
     gap             14830. .14929
                     /estimated_length=unknown
     misc_feature    14930. .17244
                     /note="assembly_name:Contig21"
     gap             17245. .17344
                     /estimated_length=unknown
     misc_feature    17345. .19950
                     /note="assembly_name:Contig22"
     gap             19951. .20050
                     /estimated_length=unknown
     misc_feature    20051. .23459
                     /note="assembly_name:Contig23"
     gap             23460. .23559
                     /estimated_length=unknown
     misc_feature    23560. .26491
                     /note="assembly_name:Contig24"
     gap             26492. .26591

  Query Match             17.8%;  Score 333.4;  DB 12;  Length 201508;
  Best Local Similarity   81.6%;  Pred. No. 1.2e-94;
  Matches  399;  Conservative    0;  Mismatches   86;  Indels    4;  Gaps    1;

Qy        502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
              ||| ||| |||  || ||  ||| ||||||||||||||||  |||||||  ||||||||||
Db     179381 TTTATTTATTTATTTATTTATTTATTTTGAGACAGAGTCTCACTCTGTCTCCCAGGCTGG 179440

Qy        562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
              ||| ||||||||||||||  || ||||||||||||||||||| | | ||||||  ||||||
Db     179441 AGTTCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCTCAGGTTCAAGTGATTCTT 179500

Qy        622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----TT 677
              |||||||||||||||  ||||||||  ||||||||  |  | ||||||||||  ||    ||
Db     179501 CTGCCTCAGCCTCCTGAGTAGCTGCAATTACAGGCGTGCACCACCACACCTGGCTAATTT 179560

Qy        678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
              |||||  |||| ||||  ||||||  || ||||  ||||||  ||||||  |||||||  |||||
Db     179561 TTTTTTTTTTTGAGATGGAGTCTCGCTCTGTTGCCCAGGTTGGAGTGCAGTGGCGTGATC 179620

Qy        738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
              || | |||||||  ||||||||||||  ||||||||||  ||||||||||||||||  ||||  |
Db     179621 TCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCTA 179680

Qy        798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
              ||||||  |||  |||||  ||||  ||  |||  |  ||||||||||||||||||||||||||||
Db     179681 GTAGCTGGGACTACAGGCGTGCGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAG 179740

Qy        858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
              |||  |  |||  |||||||||||| |||  ||||  ||||  ||  |||||||  |||||||||
Db     179741 ATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCC 179800

Qy        918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTA 977
```

```
          |||  |||||||||||||||||||| ||| |||||||| || ||| |||||     |
   179801 ACCTCAGCCTCCCAAAGTGCTGGGATTTCAGCCGTGAGCCACCGTGCCCGGCCTATTTTT 179860

      978 CTCCTTTAG 986
          |  |||||
   179861 ATTTTTTAG 179869


;T 8
)320/c
;       AC090320              191353 bp    DNA     linear   HTG 09-MAY-2001
;ITION  Homo sapiens chromosome 18 clone RP11-892H1 map 18, WORKING DRAFT
        SEQUENCE, 10 unordered pieces.
;SION   AC090320
:ON     AC090320.4  GI:13518166
:ORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
:E      Homo sapiens (human)
;ANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
RENCE   1  (bases 1 to 191353)
:HORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
:LE     Homo sapiens chromosome 18, clone RP11-892H1
)URNAL  Unpublished
RENCE   2  (bases 1 to 191353)
:HORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
        Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
        Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
        Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
        Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
        Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
        Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
        Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
        Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
        Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
        McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
        Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
        O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
        Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
        Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
        Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
        Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
        Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
        Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
        Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
        Zembek,L., Zimmer,A. and Zody,M.
:LE     Direct Submission
)URNAL  Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
        Research, 320 Charles Street, Cambridge, MA 02141, USA
;NT     On Apr 3, 2001 this sequence version replaced gi:13470217.
        All repeats were identified using RepeatMasker:
        Smit, A.F.A. & Green, P. (1996-1997)
        http://ftp.genome.washington.edu/RM/RepeatMasker.html
        -------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
        -------------- Project Information
            Center project name: L12591
            Center clone name: 892_H_1
        -------------- Summary Statistics
            Sequencing vector: Plasmid; n/a; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 187356 bases at least Q40
            Consensus quality: 189415 bases at least Q30
            Consensus quality: 190119 bases at least Q20
            Insert size: 183000; agarose-fp
            Insert size: 190453; sum-of-contigs
            Quality coverage: 6.0 in Q20 bases; agarose-fp
            Quality coverage: 5.7 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     7710: contig of 7710 bp in length
            *     7711     7810: gap of 100 bp
            *     7811     8914: contig of 1104 bp in length
            *     8915     9014: gap of 100 bp
            *     9015    10614: contig of 1600 bp in length
            *    10615    10714: gap of 100 bp
            *    10715    14053: contig of 3339 bp in length
            *    14054    14153: gap of 100 bp
            *    14154    21518: contig of 7365 bp in length
            *    21519    21618: gap of 100 bp
            *    21619    75714: contig of 54096 bp in length
            *    75715    75814: gap of 100 bp
            *    75815    99969: contig of 24155 bp in length
            *    99970   100069: gap of 100 bp
            *   100070   130859: contig of 30790 bp in length
            *   130860   130959: gap of 100 bp
            *   130960   167261: contig of 36302 bp in length
            *   167262   167361: gap of 100 bp
            *   167362   191353: contig of 23992 bp in length.
FEATURES             Location/Qualifiers
     source          1. .191353
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="18"
                     /map="18"
                     /clone="RP11-892H1"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .7710
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     gap             7711. .7810
                     /estimated_length=100
     misc_feature    7811. .8914
                     /note="assembly_fragment"
     gap             8915. .9014
                     /estimated_length=100
     misc_feature    9015. .10614
                     /note="assembly_fragment"
     gap             10615. .10714
                     /estimated_length=100
     misc_feature    10715. .14053
                     /note="assembly_fragment"
     gap             14054. .14153
                     /estimated_length=100
     misc_feature    14154. .21518
                     /note="assembly_fragment"
     gap             21519. .21618
                     /estimated_length=100
     misc_feature    21619. .75714
                     /note="assembly_fragment"
     gap             75715. .75814
                     /estimated_length=100
     misc_feature    75815. .99969
                     /note="assembly_fragment"
     gap             99970. .100069
                     /estimated_length=100
     misc_feature    100070. .130859
                     /note="assembly_fragment"
     gap             130860. .130959
                     /estimated_length=100
     misc_feature    130960. .167261
                     /note="assembly_fragment"
```

```
  gap             167262. .167361
                  /estimated_length=100
  misc_feature    167362. .191353
                  /note="assembly_fragment
                  clone_end:T7
                  vector_side:right"
:IN

:ery Match          17.8%;  Score 333;  DB 12;  Length 191353;
:st Local Similarity  82.6%;  Pred. No. 1.5e-94;
:tches  394;  Conservative    0;  Mismatches   80;  Indels    3;  Gaps    1;

       498 CTGCTTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGG 557
           ||.||| |  ||||||| |||||| ||||||||||||||||||||||||||| | |||||||
    187611 CTGCTCTACTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTTTGCCCAGG 187552

       558 CTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGAT 617
           |||||||||||||||  |||||| || || |||||||| ||||||| | |||||| ||
    187551 CTGGAGTGCAGTGGCGCGATCTCGGCTCATTGCAACCTCCGCCTCCCAGGTTCAAGCAAT 187492

       618 TCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-- 675
           | ||||||||||||||||||| ||||||||||||||||  | ||||||||||| ||
    187491 TTTCCTGCCTCAGCCTCCCAAATAGCTGGGATTACAGGCACCTGCCACCACACCCGGCTA 187432

       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
            ||||||| |||| ||| |||||||||| |||||||| |||||| ||| |||||||| |
    187431 ATTTTTTTTTTTTCTAGATGGAGTCTTGCTCTGTCACCTAGGCTGAAGTGCAGTGGCAAG 187372

       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
           |||||  |||||| |||||||| |||||||||||||||||||||||||||||||||| ||||
    187371 ATCTCCACTCACTGCAACCTCCACATCCCGGGTTCAAGCAATTCTCCTGCCTCAGGCTCC 187312

       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
           || |||||| |||||||||        || ||||  |  ||||||||||||||| ||| |
    187311 TGTGTAGCTGGGATTACAGGCACCTGCCACCACACTTGGCTAATTTTTGTATTTGTAGCA 187252

       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
           |||||| | |||  |||||||||||| ||| | |  ||| || ||||||||||||||
    187251 GAGATGAGGTTTCACCATGTTGGCCAGGCTGCTTTCAAACTCCTGACCTCAGGTGATCTA 187192

       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
           |   |||||||||| |||||||||| ||| |||  |||||||| | |||| |||||
    187191 CCCTCCTTGGCCTCTCAAAGTGCTGGAATTACAGATGTGAGCCACCACGCCTGGCCT 187135


:LT 9
:3189
:S       AC113189              184349 bp    DNA     linear   PRI 04-OCT-2002
:NITION  Homo sapiens chromosome 17, clone RP11-104H15, complete sequence.
:SSION   AC113189
:ION     AC113189.11  GI:23499591
:ORDS    HTG.
:CE      Homo sapiens (human)
:GANISM  Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
         Hominidae; Homo.
:ERENCE  1  (bases 1 to 184349)
:THORS   Birren,B., Nusbaum,C. and Lander,E.
:TLE     Homo sapiens chromosome 17, clone RP11-104H15
:URNAL   Unpublished
:ERENCE  2  (bases 1 to 184349)
:THORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
         Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
         Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
         Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
         Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
         Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
         Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
         Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
         Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
         Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
         Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
         McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
         Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
         Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
         Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
         Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
         Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
         Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
         Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
         Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
         Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
         Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (26-FEB-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3  (bases 1 to 184349)
  AUTHORS    Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
             Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
             Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
             Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
             Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
             Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
             Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
             Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
             Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
             Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
             Zembek,L., Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (18-SEP-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    4  (bases 1 to 184349)
  AUTHORS    Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
             Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
             Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
             Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
             Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
             Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
             Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
             Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
             Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
             Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
             Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
             Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
             Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
             Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
             Zembek,L., Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (04-OCT-2002) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Oct 4, 2002 this sequence version replaced gi:23129328.
             All repeats were identified using RepeatMasker:
             Smit, A.F.A. & Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             -------------- Genome Center
                 Center: Whitehead Institute/ MIT Center for Genome Research
                 Center code: WIBR
                 Web site: http://www-seq.wi.mit.edu
                 Contact: sequence_submissions@genome.wi.mit.edu
             -------------- Project Information
                 Center project name: L23113
                 Center clone name: 104_H_15
             --------------.
FEATURES                 Location/Qualifiers
     source              1. .184349
```

```
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-104H15"
                     /clone_lib="RPCI-11 Human Male BAC"
     unsure          182. .231
                     /note="single clone coverage"
     repeat_region   complement(667. .965)
                     /rpt_family="AluSc"
     repeat_region   complement(1147. .1227)
                     /rpt_family="L2"
     repeat_region   2125. .2168
                     /rpt_family="GC_rich"
     repeat_region   4585. .4625
                     /rpt_family="(TG)n"
     repeat_region   7353. .7544
                     /rpt_family="MER20"
     repeat_region   7552. .7671
                     /rpt_family="GA-rich"
     repeat_region   complement(7888. .7995)
                     /rpt_family="MER5B"
     repeat_region   complement(8099. .8409)
                     /rpt_family="AluSq"
     repeat_region   8421. .8456
                     /rpt_family="AT_rich"
     repeat_region   complement(8457. .8753)
                     /rpt_family="AluSx"
     unsure          8489. .8496
                     /note="<30 qual SNGL region"
     repeat_region   8754. .8787
                     /rpt_family="AT_rich"
     repeat_region   complement(8788. .8907)
                     /rpt_family="FLAM_C"
     repeat_region   complement(8935. .9026)
                     /rpt_family="MIR3"
     repeat_region   9218. .9273
                     /rpt_family="(TTCC)n"
     repeat_region   complement(9274. .9571)
                     /rpt_family="AluSx"
     unsure          9469. .9476
                     /note="<30 qual SNGL region"
     unsure          9482. .9486
                     /note="<30 qual SNGL region"
     unsure          9501. .9560
                     /note="<30 qual SNGL region"
     unsure          9587. .9595
                     /note="<30 qual SNGL region"
     unsure          9618. .9623
                     /note="<30 qual SNGL region"
     repeat_region   9777. .9884
                     /rpt_family="MIR"
     repeat_region   complement(9885. .10211)
                     /rpt_family="AluSx"
     repeat_region   10212. .10228
                     /rpt_family="MIR"
     repeat_region   10237. .10378
                     /rpt_family="MER104"
     repeat_region   10379. .10680
                     /rpt_family="AluSx"
     repeat_region   10681. .10735
                     /rpt_family="AT_rich"
     repeat_region   10818. .10967
                     /rpt_family="AluJo"
     repeat_region   10968. .11265
                     /rpt_family="AluSg"
     repeat_region   11266. .11442
                     /rpt_family="AluJo"
     repeat_region   11602. .11902
                     /rpt_family="AluSx"
     repeat_region   complement(11996. .12227)
                     /rpt_family="L2"
     repeat_region   complement(12220. .12452)
                     /rpt_family="L2"
     repeat_region   12544. .12842
                     /rpt_family="AluSc"
     repeat_region   complement(12886. .13197)
                     /rpt_family="AluSx"
     repeat_region   complement(13275. .13576)
                     /rpt_family="AluJb"
     repeat_region   13683. .14116
                     /rpt_family="MLT1G3"
     repeat_region   complement(14117. .14424)
                     /rpt_family="AluY"
     repeat_region   14425. .14454
                     /rpt_family="MLT1G3"
     repeat_region   14509. .14632
                     /rpt_family="MIR"
     repeat_region   complement(14674. .14962)
                     /rpt_family="AluSx"
     repeat_region   complement(14963. .15131)
                     /rpt_family="MLT1F"
     repeat_region   complement(15137. .15248)
                     /rpt_family="MLT1F"
     repeat_region   complement(15249. .15437)
                     /rpt_family="MIR"
     repeat_region   15640. .15933
                     /rpt_family="AluY"
     repeat_region   complement(16136. .16283)
                     /rpt_family="L2"
```

Query Match 17.8%; Score 332.6; DB 5; Length 184349;
Best Local Similarity 80.3%; Pred. No. 2.1e-94;
Matches 404; Conservative 0; Mismatches 94; Indels 5; Gaps 1;

```
Qy        500 GCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCT 559
              | | |||||||| | ||  | | ||||||||||||||||| ||  ||| |||  |||||||
Db      16366 GGTGTTTTTTTTGTTTTGATTTGTTTTTTTGAGACAGAGCCTCACTCCGTCATCCAGGCT 16425

Qy        560 GGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTC 619
              ||| |||||||||| |||||| || ||||||||| ||||||| || |||||||| |||||
Db      16426 GGATTGCAGTGGCACGATCTCGGCTCACTGCAACTTCTGCCTTCCAGATTCAAGTGATTC 16485

Qy        620 TCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---- 675
              |||||||||||||||| |||||||||||||||||| || ||||||||  ||||||
Db      16486 TCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCGCCACCATGCCCAGCTAAT 16545

Qy        676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
               ||||||| |||| ||||  ||||||||| |||||||||||||||||||| |||||||  |
Db      16546 TTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCG 16605

Qy        735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
              ||||| | ||||||| |||||  |||||  ||||| |||| |||||||||| |||| ||||
Db      16606 ATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTTAAGCGATTCTCCTGCTTCAGCCTCC 16665

Qy        795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
              ||||||||  ||||||||      | || ||| |  ||||||||||||||||||||||||
Db      16666 AGAGTAGCTGCGATTACAGGCACACGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTA 16725

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              |||| | | |||  |||||||||||| ||| |||| |||| || ||||||||||||||||
Db      16726 GAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCA 16785

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
              |   |||  |||||| |||||||||||||| ||||||||||||  ||||| |||| | |
Db      16786 CCCACCTCAGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCGGCCAGGG 16845

Qy        975 CTACTCCTTTAGTCTCTGGAAAG 997
              ||    | | |  | | | |||
Db      16846 ATAGGTGTCTTGCATTTAGCAAG 16868
```

RESULT 10
CNS01DVI/c
LOCUS CNS01DVI 187710 bp DNA linear PRI 04-OCT-2001

```
:NITION  Human chromosome 14 DNA sequence BAC R-362L22 of library RPCI-11
         from chromosome 14 of Homo sapiens (Human), complete sequence.
:SSION   AL135838
:ION     AL135838.6  GI:15982201
:ORDS    HTG.
:CE      Homo sapiens (human)
:GANISM  Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
         Hominidae; Homo.
:RENCE   1  (bases 1 to 187710)
:THORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
         Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
         Levy,M., Eckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
         Gyapay,G., Saurin,W. and Weissenbach,J.
:TLE     Sequencing of the human chromosome 14
:URNAL   Unpublished
:RENCE   2  (bases 1 to 187710)
:THORS   Genoscope.
:TLE     Direct Submission
:URNAL   Submitted (04-OCT-2001) Genoscope - Centre National de Sequencage :
         BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
         - Web : www.genoscope.cns.fr)
:ENT     On Oct 5, 2001 this sequence version replaced gi:13872725.
         --------------- Genome Center
         Center: Genoscope / Centre National de Sequencage
         Center code: GS
         Web site: http://www.genoscope.cns.fr/
         Contact: SeqRef@genoscope.cns.fr
         ---------------
         The following BAC sequence is oriented from the T7 to the SP6 end.
         Upstream BAC (overlapping the T7 end) : C-2644I21 (AC=AL163974)
         Downstream BAC (overlapping the SP6 end) : R-638I2 ---------------
         Summary Statistics
         Assembly program: Phrap; version 2.0
         Quality coverage: 6.92x in Q20 bases; sum-of-contigs
         ---------------
         ------------------------
         Overall quality chart :
         Range     :   bases
         0         :       1
         1  -  9   :      53
         10 - 19   :     407
         20 - 29   :    1172
         30 - 39   :    5394
         40 - 49   :   11391
         50 - 59   :   10988
         60 - 69   :   11109
         70 - 79   :   20167
         80 - 89   :   49419
         90 - 99   :   77609
         ------------------------
         Percentage of bases with a quality value >= 40 : 96 %.
:URES              Location/Qualifiers
  source           1. .187710
                   /organism="Homo sapiens"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:9606"
                   /chromosome="14"
                   /clone="R-362L22"
                   /clone_lib="RPCI-11"
  STS              72840. .72970
                   /note="matching EMBL:R17876
                   RHdb:RH26378
                   dbSTS:STS8590
                   Identified using the e-PCR software (G. Schuler)"
  STS              87948. .88130
                   /note="matching EMBL:AA251131
                   RHdb:RH103649
                   dbSTS:STS71178
                   Identified using the e-PCR software (G. Schuler)"
  STS              111707. .111823
                   /note="matching EMBL:H13689
                   RHdb:RH44843
                   dbSTS:STS37911
                   Identified using the e-PCR software (G. Schuler)"
  STS              116864. .117047
                   /note="matching EMBL:AA085408
                   RHdb:RH65341
                   dbSTS:STS45273
                   Identified using the e-PCR software (G. Schuler)"
  STS              119261. .119399
                   /note="matching EMBL:X59892
                   RHdb:RH70848
                   dbSTS:STS50705
                   Identified using the e-PCR software (G. Schuler)"
  STS              119304. .119430
                   /note="matching EMBL:T63858
                   RHdb:RH53514
                   dbSTS:STS32482
                   Identified using the e-PCR software (G. Schuler)"
  STS              119989. .120114
                   /note="matching EMBL:H28114
                   RHdb:RH53813
                   dbSTS:STS26412
                   Identified using the e-PCR software (G. Schuler)"
  STS              162612. .162786
                   /note="matching EMBL:H37767
                   RHdb:RH77775
                   RHdb:RH15727
                   dbSTS:STS13361
                   Identified using the e-PCR software (G. Schuler)"
  STS              171231. .171463
                   /note="matching EMBL:D11738
                   RHdb:RH53906
                   dbSTS:STS42918
                   Identified using the e-PCR software (G. Schuler)"
  STS              174629. .174757
                   /note="matching EMBL:T78109
                   RHdb:RH25853
                   dbSTS:STS24160
                   Identified using the e-PCR software (G. Schuler)"
  STS              174966. .175173
                   /note="matching EMBL:H38818
                   RHdb:RH68412
                   dbSTS:STS48319
                   Identified using the e-PCR software (G. Schuler)"
  STS              176200. .176323
                   /note="matching EMBL:M77698
                   RHdb:RH44091
                   RHdb:RH1618
                   dbSTS:STS16477
                   Identified using the e-PCR software (G. Schuler)"
  STS              186933. .187058
                   /note="matching EMBL:G11457
                   RHdb:RH1924
                   dbSTS:STS19283
                   Identified using the e-PCR software (G. Schuler)"
ORIGIN

  Query Match             17.8%;  Score 332.6;  DB 5;  Length 187710;
  Best Local Similarity   82.6%;  Pred. No. 2.1e-94;
  Matches  394;  Conservative    0;  Mismatches   79;  Indels    4;  Gaps    1;

Qy        499 TGCTTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGC 558
              | || ||  |  |||||||||| ||||||||||||||||||||||||||||| ||||||||
Db      22393 TTCTGTTGGTAATTTCTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTTGCCCAGGC 22334

Qy        559 TGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATT 618
              ||||||||| ||| | |||||||||| |||||||||||| | |||  || ||||||  ||||
Db      22333 TGGAGTGCAATGGGACGATCTCTGCTCACTGCAACCTCCGACTCTGGGGTTCAAGTGATT 22274

Qy        619 CTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC--- 675
              ||||||||| ||||||||||||||||||||||||||||||  |||||||  |||| |
Db      22273 CTCCTGCCTTAGCCTCCCAAGTAGCTGGGATTACAGGTGCCTGCCACCATGCCCAACTAA 22214
```

```
     676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
          ||||||| |||| ||||  |||||||| ||||  |||||||||||||| ||||||||||
   22213 TTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCATG 22154

     735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
          ||||| | ||||||| |||||  ||||||  |||||||  ||||||  |||||||| ||||
   22153 ATCTCGGCTCACTGCAACCTCTGCCTCCCAAGTTCAAGTGATTCTCTTGCCTCAGCCTCC 22094

     795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
           |||||||||||||||||  |  | || ||| |  |||||||||||||||||||||||||
   22093 CAAGTAGCTAGGATTACAGGTGCCCGCCACCAAGCCCAGCTAATTTTTGTATTTTTAGTA 22034

     855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
          |||||| | |||  |||  |||  || ||| |||  ||||  | ||||||||||||||||
   22033 GAGATGGGGTTTCACCAGATTGATCAGGCTGGTCCCGAACTACTGACCTCAGGTGATCCA 21974

     915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
          | ||||  |||||||||||||||||||||| |||||||||||||| ||| ||  ||||
   21973 CCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCCAGCCT 21917
```

```
.LT 11
5375
.S       AC145375              217260 bp    DNA     linear   PRI 27-APR-2005
'INITION  Pan troglodytes BAC clone RP43-57M5 from chromosome 7, complete
          sequence.
.SSION    AC145375
:ON       AC145375.5  GI:60301748
)RDS      HTG.
:E        Pan troglodytes (chimpanzee)
GANISM    Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
          Hominidae; Pan.
RENCE     1  (bases 1 to 217260)
'HORS     Shah,N., Kozlowicz,A. and Haglund,K.
:LE       The sequence of Pan troglodytes BAC clone RP43-57M5
JRNAL     Unpublished (2001)
RENCE     2  (bases 1 to 217260)
'HORS     Wilson,R.K.
:LE       Direct Submission
JRNAL     Submitted (28-JUN-2003) Genetics, Genome Sequencing Center, 4444
          Forest Park Parkway, St. Louis, MO 63108, USA
RENCE     3  (bases 1 to 217260)
'HORS     Wilson,R.K.
:LE       Direct Submission
JRNAL     Submitted (14-JAN-2005) Genetics, Genome Sequencing Center, 4444
          Forest Park Parkway, St. Louis, MO 63108, USA
RENCE     4  (bases 1 to 217260)
'HORS     Wilson,R.K.
:LE       Direct Submission
JRNAL     Submitted (26-FEB-2005) Genetics, Genome Sequencing Center, 4444
          Forest Park Parkway, St. Louis, MO 63108, USA
RENCE     5  (bases 1 to 217260)
'HORS     Wilson,R.K.
:LE       Direct Submission
JRNAL     Submitted (27-APR-2005) Washington University School of Medicine,
          Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO
          63108, USA
'NT       On Feb 26, 2005 this sequence version replaced gi:57790261.
          -------------- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu
              Contact: submissions@watson.wustl.edu
          -------------- Summary Statistics
              Center project name: C_PT057M05


          NOTICE:

          This sequence was finished as follows unless otherwise noted:
          all regions were double stranded, sequenced with an alternate
          chemistry, or covered by high quality data (i.e., phred quality >=
          30); an attempt was made to resolve all sequencing problems, such
          as compressions and repeats; all regions were covered by sequence
          from more than one subclone; and the assembly was confirmed by
          restriction digest.

          MAPPING INFORMATION:
          Mapping information for this clone was provided by Dr. Wes Warren,
          Department of Genetics, Washington University, St. Louis MO. For
          additional information about the map position of this sequence, see
          http://genome.wustl.edu

          SOURCE INFORMATION:
          The RPCI-43 BAC Library has been constructed by Chung-Li Shu. DNA
          was isolated from white blood cells obtained from a male chimpanzee
          (Pan troglodytes,'Clint', Yerkes #C0471; birthdate: 6-6-80). The
          clone and detailed information can be obtained from ResGen
          (http://www.resgen.com) or Pieter de Jong and co-workers at
          http://www.bacpac.chori.org.

          This sequence is the entire insert of the clone.
FEATURES             Location/Qualifiers
     source          1. .217260
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /chromosome="7"
                     /clone="RP43-57M5"
                     /clone_lib="RPCI-43"
     unsure          21655. .21761
                     /note="Sequence derived from one plasmid subclone."
     unsure          58557
                     /note="Unresolved bases"
     unsure          63511. .63603
                     /note="Sequence derived from one plasmid subclone."
     unsure          96234. .96247
                     /note="Sequence derived from one plasmid subclone."
     misc_feature    97063. .97100
                     /note="Sequence derived from PCR product of project DNA"
     unsure          213623. .213674
                     /note="Sequence derived from one plasmid subclone."
ORIGIN

  Query Match             17.8%;  Score 332;  DB 5;  Length 217260;
  Best Local Similarity   81.8%;  Pred. No. 3.3e-94;
  Matches  383;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy        503 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 562
              ||| ||| ||| || || || || ||||||||||||||||||||||||||||||||||||
Db      77450 TTTATTTATTTATTGTTATTATTATTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 77509

Qy        563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
              |||||||| ||||||||  || ||||||||||||  | ||| || || |||| |||| ||
Db      77510 GTGCAGTGTCATGATCTTGGCTCACTGCAACCTCCACTTCCTGGGTTTAAGCAATTCCCC 77569

Qy        623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTT 682
              || ||||||| ||||| |||||||||||||||| |   |||||||  ||| |||||||||
Db      77570 TGTCTCAGCCACCCAAATAGCTGGGATTACAGGCGTGTGCCACCATGCCCGGCTTTTTTT 77629

Qy        683 ATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGT 742
              |||| |||   ||||||||| ||||||| |||||||||||  ||| |||| |||||  |
Db      77630 TTTTTTGAGGTGGAGTCTTGCTCTGTCACTCAGGCTGGAGAGCAGCGGCACGATCTTGGC 77689

Qy        743 TCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGC 802
              ||||||| |||||| ||| |||||||||| ||||||| || ||||| |||| ||||||||
Db      77690 TCACTGCAACCTCCGCCTACCGGGTTCAAGTAATTCTCGTGTCTCAGCCTCCCGAGTAGC 77749

Qy        803 TAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCG 862
              | ||| |||||  |  |||| ||||     ||||||||||||||||||||||||  | |
Db      77750 TGGGACTACAGGTGCCCACCACCACACCTGGCTAATTTTTGTATTTTTAGTAGAGCCGGG 77809

Qy        863 CTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTT 922
              ||| ||||||||||||  ||| |||| ||   || || ||||||||||||||  |||||
```

```
     77810 GTTTCGCCATGTTGGCCGGGCTGGTCTCGATATCCTGATCTCAGGTGATCCGCCCGCCTT 77869

       923 GGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
           |||||||||||||| ||||||| |||||||||||||   |||||  |||
     77870 GGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCCAATGCGCCTAGCC 77917


LT 12
0202_09/c
MMENT
quence split into 10 fragments  LOCUS AL390202 Accession AL390202
  Fragment Name      Begin     End
  AL390202_00            1    110000
  AL390202_01       100001    210000
  AL390202_02       200001    310000
  AL390202_03       300001    410000
  AL390202_04       400001    510000
  AL390202_05       500001    610000
  AL390202_06       600001    710000
  AL390202_07       700001    810000
  AL390202_08       800001    910000
  AL390202_09       900001    988176
inuation (10 of 10) of AL390202 from base 900001 (AL390202 Homo sapiens chromosome 2

ery Match             17.7%;  Score 331.2;  DB 12;  Length 88176;
st Local Similarity   80.0%;  Pred. No. 4.8e-94;
tches  405;  Conservative    0;  Mismatches   93;  Indels    8;  Gaps    1;

       500 GCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCT 559
           | | |||||||| | ||  | | ||||||||||||||||| ||  ||| |||  |||||||
     33917 GGTGTTTTTTTTGTTTTGATTTGTTTTTTTTGAGACAGAGCCTCACTCCGTCATCCAGGCT 33858

       560 GGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTC 619
           ||| |||||||||| |||||  || |||||||| ||||||| || |||||||| |||||
     33857 GGATTGCAGTGGCACGATCTTGGCTCACTGCAACTTCTGCCTTCCAGATTCAAGTGATTC 33798

       620 TCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---- 675
           ||||||||||||||||  |||||||||||||||||| || ||||||||  ||||||
     33797 TCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCGCCACCATGCCCAGCTAAT 33738

       676 ----TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGC 731
               ||||||| |||| ||||  |||||||||| |||||||||||||||||||| |||||||
     33737 TTTTTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGC 33678

       732 ATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTC 791
             |||||| | ||||||| |||||  |||||| ||||| |||| |||||||||| |||| |
     33677 GCGATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTTAAGCGATTCTCCTGCTTCAGCC 33618

       792 TCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTA 851
           ||| |||||||| |||||||||      | || ||| |  ||||||||||||||||||||
     33617 TCCAGAGTAGCTGGGATTACAGGCACACGCCACCATGCCCAGCTAATTTTTGTATTTTTA 33558

       852 GTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGAT 911
           ||||||| | | |||  |||||||||||| ||| |||||||| || ||||||||||||
     33557 GTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTGGAACTCCTGACCTCAGGTGAT 33498

       912 CCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
           || |   |||  |||||| |||||||||||||| ||||||||||||   ||||| ||||
     33497 CCACCCACCTCAGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCGGCCA 33438

       972 GAGCTACTCCTTTAGTCTCTGGAAAG 997
           | | ||    | | |  | | | |||
     33437 GGGATAGGTGTCTTGCATTTAGCAAG 33412


LT 13
32732
S       AL162732             185353 bp    DNA     linear   PRI 14-NOV-2001
NITION  Human DNA sequence from clone RP11-32M23 on chromosome 9, complete
        sequence.
SSION   AL162732
ION     AL162732.43  GI:16516845
ORDS    HTG.

SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 185353)
  AUTHORS   Sehra,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 30, 2001 this sequence version replaced gi:16508195.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr9
            RP11-32M23 is from the library RPCI-11.1 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            This sequence is the entire insert of clone RP11-32M23 The true
            left end of clone RP11-576C12 is at 99588 in this sequence. The
            true right end of clone RP11-432F4 is at 14963 in this sequence.
FEATURES              Location/Qualifiers
     source           1. .185353
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="9"
                      /clone="RP11-32M23"
                      /clone_lib="RPCI-11.1"
     misc_feature     45498. .45519
                      /note="Single clone region. Sequence from reads from a
                      short insert library derived from a single pUC clone.
                      Restriction digest data confirm the assembly."
     misc_feature     135906. .136041
                      /note="Single clone region. Sequence from reads from a
                      short insert library derived from a single pUC clone.
                      Restriction digest data confirm the assembly."
     misc_feature     136295. .136325
                      /note="Single clone region. Sequence from reads from a
                      short insert library derived from a single pUC clone.
                      Restriction digest data confirm the assembly."
     misc_feature     136675. .136803
                      /note="Single clone region. Sequence from reads from a
                      short insert library derived from a single pUC clone.
                      Restriction digest data confirm the assembly."
ORIGIN

  Query Match             17.7%;  Score 331.2;  DB 5;  Length 185353;
  Best Local Similarity   80.0%;  Pred. No. 5.8e-94;
  Matches  405;  Conservative    0;  Mismatches   93;  Indels    8;  Gaps    1;

Qy        500 GCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCT 559
              | | |||||||| | ||  | | ||||||||||||||||| ||  ||| |||  |||||||
```

```
134283 GGTGTTTTTTTTGTTTTGATTTGTTTTTTTGAGACAGAGCCTCACTCCGTCATCCAGGCT 134342

   560 GGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTC 619
       ||| |||||||||| |||||  || |||||||| ||||||| || |||||||| |||||
134343 GGATTGCAGTGGCACGATCTTGGCTCACTGCAACTTCTGCCTTCCAGATTCAAGTGATTC 134402

   620 TCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---- 675
       |||||||||||||||  |||||||||||||||||| || ||||||||  ||||||
134403 TCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCGCCACCATGCCCAGCTAAT 134462

   676 ----TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGC 731
           ||||||| |||| |||| |||||||||| |||||||||||||||||||| |||||||
134463 TTTTTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGC 134522

   732 ATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTC 791
         |||||| | ||||||| |||||  ||||| ||||| |||| |||||||||| |||| |
134523 GCGATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTTAAGCGATTCTCCTGCTTCAGCC 134582

   792 TCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTA 851
       ||| ||||||||| |||||||||     | || ||| |  ||||||||||||||||||||
134583 TCCAGAGTAGCTGGGATTACAGGCACACGCCACCATGCCCAGCTAATTTTTGTATTTTTA 134642

   852 GTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGAT 911
       ||||||| | | |||  |||||||||||| ||| |||||||| || ||||||||||||
134643 GTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTGGAACTCCTGACCTCAGGTGAT 134702

   912 CCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
       || |   |||  |||||| |||||||||||||| |||||||||||||   ||||| ||||
134703 CCACCCACCTCAGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCGGCCA 134762

   972 GAGCTACTCCTTTAGTCTCTGGAAAG 997
       | | ||   | | |  | | | |||
134763 GGGATAGGTGTCTTGCATTTAGCAAG 134788
```

```
LT 14
5954
S       AC026954              101029 bp    DNA     linear   PRI 10-SEP-2002
NITION  Homo sapiens chromosome 17, clone RP11-542C16, complete sequence.
SSION   AC026954
ON      AC026954.14  GI:22773348
RDS     HTG.
E       Homo sapiens (human)
ANISM   Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
ENCE    1  (bases 1 to 101029)
HORS    Birren,B., Nusbaum,C. and Lander,E.
LE      Homo sapiens chromosome 17, clone RP11-542C16
RNAL    Unpublished
ENCE    2  (bases 1 to 101029)
HORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
        Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
        Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
        Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
        Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
        Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
        Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
        Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
        Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
        Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
        Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
        McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
        Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
        Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
        O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
        Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
        Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
        Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
        Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
        Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
        Young,G., Zainoun,J., Zimmer,A. and Zody,M.
```

```
TITLE     Direct Submission
JOURNAL   Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3  (bases 1 to 101029)
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
          Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
          Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
          Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
          Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
          Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
          Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
          Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
          McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
          O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
          Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
          Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
          Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
          Zembek,L., Zimmer,A. and Zody,M.
TITLE     Direct Submission
JOURNAL   Submitted (22-AUG-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4  (bases 1 to 101029)
AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
          Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
          Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
          Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
          Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
          Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
          Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
          Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
          Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
          McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
          Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
          O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
          Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
          Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
          Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
          Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
          Zembek,L., Zimmer,A. and Zody,M.
TITLE     Direct Submission
JOURNAL   Submitted (10-SEP-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT   On Sep 10, 2002 this sequence version replaced gi:22417404.
          All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          -------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
          -------------- Project Information
              Center project name: L8435
              Center clone name: 542_C_16
          --------------

          Only the last 101.0 kilobases of this clone are being submitted.
          The remainder overlaps accession number AC003688 [WICGR project
          L270].
FEATURES             Location/Qualifiers
     source          1. .101029
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-542C16"
                     /clone_lib="RPCI-11 Human Male BAC"
```

```
repeat_region   complement(1. .70)
                /rpt_family="AluS"
repeat_region   163. .463
                /rpt_family="AluSx"
repeat_region   560. .710
                /rpt_family="AluJb"
repeat_region   713. .1021
                /rpt_family="AluSp"
repeat_region   complement(1196. .1491)
                /rpt_family="AluSx"
repeat_region   complement(1811. .2028)
                /rpt_family="L2"
repeat_region   complement(4307. .4418)
                /rpt_family="MIR"
repeat_region   8346. .8391
                /rpt_family="C-rich"
repeat_region   complement(8591. .8895)
                /rpt_family="AluY"
repeat_region   complement(10340. .10628)
                /rpt_family="AluSx"
repeat_region   complement(13036. .13210)
                /rpt_family="MER5A"
repeat_region   15831. .16010
                /rpt_family="(CCG)n"
repeat_region   16218. .16245
                /rpt_family="GC_rich"
repeat_region   complement(16726. .16853)
                /rpt_family="AluSg/x"
repeat_region   complement(16950. .17263)
                /rpt_family="AluSx"
repeat_region   17389. .17481
                /rpt_family="L2"
repeat_region   17512. .17699
                /rpt_family="L2"
repeat_region   17772. .17820
                /rpt_family="Alu"
repeat_region   18497. .18810
                /rpt_family="AluSq"
repeat_region   18825. .18961
                /rpt_family="AluSx"
repeat_region   18962. .19269
                /rpt_family="AluY"
repeat_region   19270. .19443
                /rpt_family="AluSx"
repeat_region   19446. .19740
                /rpt_family="AluY"
repeat_region   complement(19936. .20241)
                /rpt_family="AluY"
repeat_region   complement(20242. .20547)
                /rpt_family="AluJb"
repeat_region   20716. .21015
                /rpt_family="AluJo"
repeat_region   21018. .21327
                /rpt_family="AluSq"
repeat_region   complement(21665. .21963)
                /rpt_family="AluSx"
repeat_region   22317. .22624
                /rpt_family="AluJo"
unsure          complement(22342. .22348)
                /note="<30 qual SNGL region"
unsure          complement(22373. .22434)
                /note="single clone coverage"
unsure          complement(22431. .22437)
                /note="<30 qual SNGL region"
unsure          complement(22434)
                /note="Probably A; possibly C"
unsure          complement(22479. .22500)
                /note="<30 qual SNGL region"
unsure          complement(22512. .22516)
                /note="<30 qual SNGL region"
repeat_region   complement(22811. .23148)
                /rpt_family="AluY"
repeat_region   complement(23171. .23454)
                /rpt_family="AluSq"
repeat_region   complement(23811. .24108)
                /rpt_family="AluJb"
unsure          24029. .24061
                /note="PCR product sequence only"
repeat_region   complement(24158. .24212)
                /rpt_family="AluSp"
repeat_region   complement(24216. .24528)
                /rpt_family="AluSg"
repeat_region   complement(24552. .24863)
                /rpt_family="AluY"
repeat_region   24928. .25220
                /rpt_family="AluSg"
unsure          25064. .25069
                /note="<30 qual SNGL region"
unsure          25117. .25124
                /note="<30 qual SNGL region"
repeat_region   complement(25261. .25436)
```

```
Query Match             17.7%;  Score 331;  DB 5;  Length 101029;
Best Local Similarity   80.1%;  Pred. No. 5.8e-94;
Matches  403;  Conservative    0;  Mismatches   95;  Indels    5;  Gaps    1;

Qy        500 GCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCT 559
              | | ||||||||| | ||  | | ||||||||||||||||||| ||  ||| |||  |||||||
Db      87829 GGTGTTTTTTTTGTTTTGATTTGTTTTTTTGAGACAGAGCCTCACTCCGTCATCCAGGCT 87888

Qy        560 GGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTC 619
              ||| |||||||||| |||||| || ||||||||| ||||||| || |||||||| |||||
Db      87889 GGATTGCAGTGGCACGATCTCGGCTCACTGCAACTTCTGCCTTCCAGATTCAAGTGATTC 87948

Qy        620 TCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---- 675
              |||||||||||||||||  || ||||||||||||||| || |||||||||  ||||||
Db      87949 TCCTGCCTCAGCCTCCTGAGAAGCTGGGATTACAGGCGCCCGCCACCATGCCCAGCTAAT 88008

Qy        676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
               ||||||| |||| ||||  |||||||| |||||||||||||||||||| |||||||  |
Db      88009 TTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCG 88068

Qy        735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
              ||||| | ||||||| |||||  ||||| ||||| |||| |||||||||| |||| ||||
Db      88069 ATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTTAAGCGATTCTCCTGCTTCAGCCTCC 88128

Qy        795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
              ||||||||  ||||||||      | || ||| |  ||||||||||||||||||||||||
Db      88129 AGAGTAGCTGCGATTACAGGCACACGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTA 88188

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              |||| | | |||  |||||||||||| ||| |||| |||| || ||||||||||||||||
Db      88189 GAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCA 88248

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
              |   |||  |||||| |||||||||||||| ||||||||||||  ||||| |||| | |
Db      88249 CCCACCTCAGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCGGCCAGGG 88308

Qy        975 CTACTCCTTTAGTCTCTGGAAAG 997
              ||    | | |  | | | |||
Db      88309 ATAGGTGTCTTGCATTTAGCAAG 88331
```

```
RESULT 15
AC087388/c
LOCUS       AC087388               121017 bp    DNA     linear   PRI 08-OCT-2002
DEFINITION  Homo sapiens chromosome 17, clone RP11-199F11, complete sequence.
ACCESSION   AC087388
VERSION     AC087388.9  GI:23592178
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 121017)
```

```
:HORS   Birren,B., Nusbaum,C. and Lander,E.
:LE     Homo sapiens chromosome 17, clone RP11-199F11
JRNAL   Unpublished
:ENCE   2  (bases 1 to 121017)
:HORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
        Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B., Brown,A.,
        Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
        Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
        Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
        Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
        Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
        Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
        Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
        Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
        McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
        Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
        O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
        Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
        Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
        Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
        Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
        Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
        Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
        Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
        Zembek,L., Zimmer,A. and Zody,M.
:LE     Direct Submission
JRNAL   Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome
        Research, 320 Charles Street, Cambridge, MA 02141, USA
:ENCE   3  (bases 1 to 121017)
:HORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
        Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
        Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
        Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
        Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
        Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
        Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
        Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
        Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
        McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
        Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
        O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
        Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
        Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
        Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
        Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
        Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
        Zembek,L., Zimmer,A. and Zody,M.
:LE     Direct Submission
:RNAL   Submitted (03-OCT-2002) Whitehead Institute/MIT Center for Genome
        Research, 320 Charles Street, Cambridge, MA 02141, USA
:ENCE   4  (bases 1 to 121017)
:HORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
        Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
        Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
        Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
        Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
        Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
        Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
        Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
        Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
        McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
        Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
        O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
        Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
        Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
        Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
        Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
        Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
        Zembek,L., Zimmer,A. and Zody,M.
:LE     Direct Submission
JRNAL   Submitted (08-OCT-2002) Whitehead Institute/MIT Center for Genome
        Research, 320 Charles Street, Cambridge, MA 02141, USA
:NT     On Oct 8, 2002 this sequence version replaced gi:22597626.
```

```
        All repeats were identified using RepeatMasker:
        Smit, A.F.A. & Green, P. (1996-1997)
        http://ftp.genome.washington.edu/RM/RepeatMasker.html
        -------------- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
        -------------- Project Information
            Center project name: L11969
            Center clone name: 199_F_11
        --------------

        Only the first 121.0 kilobases of this clone are being submitted.
        The remainder overlaps accession number AC007421 [WICGR project
        L529].
FEATURES             Location/Qualifiers
     source          1. .121017
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="17"
                     /map="17"
                     /clone="RP11-199F11"
                     /clone_lib="RPCI-11 Human Male BAC"
     unsure          628. .635
                     /note="<30 qual SNGL region"
     repeat_region   798. .1141
                     /rpt_family="L2"
     unsure          1028. .1033
                     /note="<30 qual SNGL region"
     repeat_region   complement(1142. .1443)
                     /rpt_family="AluSx"
     repeat_region   1444. .1969
                     /rpt_family="L2"
     repeat_region   complement(2081. .2250)
                     /rpt_family="Charlie8"
     repeat_region   complement(2386. .2681)
                     /rpt_family="AluY"
     repeat_region   2749. .2907
                     /rpt_family="L1M4"
     repeat_region   2915. .2949
                     /rpt_family="(TG)n"
     repeat_region   3005. .3157
                     /rpt_family="AluJb"
     repeat_region   3158. .3189
                     /rpt_family="(CA)n"
     repeat_region   3198. .3685
                     /rpt_family="L1M4"
     repeat_region   3686. .3714
                     /rpt_family="AT_rich"
     repeat_region   3718. .3993
                     /rpt_family="AluSx"
     repeat_region   4431. .4498
                     /rpt_family="HAL1"
     repeat_region   4513. .4846
                     /rpt_family="HAL1"
     repeat_region   4928. .5002
                     /rpt_family="AluS"
     repeat_region   5064. .5093
                     /rpt_family="(TAA)n"
     repeat_region   complement(5101. .5395)
                     /rpt_family="AluSg"
     repeat_region   5397. .5579
                     /rpt_family="HAL1"
     repeat_region   complement(5772. .6083)
                     /rpt_family="AluSp"
     repeat_region   6315. .6617
                     /rpt_family="AluJb"
     repeat_region   7680. .7840
                     /rpt_family="L1MC4a"
     repeat_region   7846. .7961
                     /rpt_family="FLAM_C"
```

```
  unsure          7859. .7865
                  /note="<30 qual SNGL region"
  repeat_region   complement(8097. .8400)
                  /rpt_family="AluSq"
  unsure          8289. .8293
                  /note="<30 qual SNGL region"
  unsure          8338. .8344
                  /note="<30 qual SNGL region"
  repeat_region   complement(9099. .9212)
                  /rpt_family="MIR"
  repeat_region   complement(9291. .9570)
                  /rpt_family="AluJo"
  repeat_region   complement(9585. .9746)
                  /rpt_family="L1PA13"
  repeat_region   9816. .10117
                  /rpt_family="AluSx"
  repeat_region   complement(10792. .11012)
                  /rpt_family="AluSg/x"
  repeat_region   11130. .11233
                  /rpt_family="L2"
  repeat_region   complement(11429. .11726)
                  /rpt_family="AluJb"
  repeat_region   complement(12660. .12962)
                  /rpt_family="AluSx"
  repeat_region   13148. .13453
                  /rpt_family="AluSp"
  repeat_region   13454. .13904
                  /rpt_family="AluY"
  repeat_region   complement(14783. .14991)
                  /rpt_family="MIR"
  repeat_region   complement(15683. .15986)
                  /rpt_family="AluSx"
  repeat_region   16482. .16798
                  /rpt_family="AluY"
  repeat_region   complement(18108. .18410)
                  /rpt_family="AluSq"
  repeat_region   18938. .19057
                  /rpt_family="AluJo"
  repeat_region   19058. .19361
                  /rpt_family="AluSx"
  repeat_region   19372. .19673
                  /rpt_family="AluJb"
  repeat_region   19681. .19998

ery Match            17.7%;  Score 331;  DB 5;  Length 121017;
at Local Similarity  82.1%;  Pred. No. 6e-94;
tches  395;  Conservative    0;  Mismatches   80;  Indels    6;  Gaps    1;

      502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
          | ||| |||||| ||||| ||||||||||||||| | |||  ||||||| ||||| ||||
    34111 TCTTTCTTTTTTTTTTTTTTTTTTTTTGAGACAAATTCTCACTCTGTCTCCCAGCCTGG 34052

      562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
          |||||| |||||| ||||| || ||||||||||||||||||||||  ||||| ||||||
    34051 AGTGCATTGGCATAATCTCAGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGTGATTCTT 33992

      622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
          ||||||||||||||||||||||||||||||||   || ||||||  ||||||
    33991 GTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCATACACCACCATGCCCAGCTTATTT 33932

      676 TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGA 735
          ||||| | |||| |||||||||||||  | | ||  || ||||| |||| ||||||   ||
    33931 TTTTTATTTTTTTGAGACAGAGTCTCTCTCAGTTGCCAAGGCTAGAGTGCAGTGGTGCGA 33872

      736 TCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCT 795
          |||||||||||||| |||||||||||| |||||||||| |||||||||||||| |||||
    33871 TCTCAGTTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCT 33812

      796 GAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAG 855
          |||||||| |||||||||  |||  || ||||    | ||||||||||||||||||||||
    33811 GAGTAGCTGGGATTACAGGCGTGTGCCACCACACCCGGCTAATTTTTGTATTTTTAGTAG 33752

      856 AGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGC 915
          ||||| | |||  |||||||| ||| ||| |||| |||| || | |||| |||||||
    Db  33751 AGATGGGGTTTCACCATGTTGCCCAGGCTGGTCTCGAACTCCTGTCCTCGAGTGATCCAC 33692

    Qy    916 TGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGC 975
              |||| |||||||||||||||||||||| |||||||||||| ||| ||  | ||   |
    Db  33691 CTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCCAGTCTCCCC 33632

    Qy    976 T 976
              |
    Db  33631 T 33631

    Search completed: June 22, 2006, 00:39:18
    Job time : 10376 secs
```

GenCore version 5.1.9

ucleic - nucleic search, using sw model

on: June 21, 2006, 21:42:46 ; Search time 9227 Seconds
(without alignments)
11332.952 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ence: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ng table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ched: 48236798 seqs, 27959665780 residues

number of hits satisfying chosen parameters: 96473596

num DB seq length: 0
num DB seq length: 2000000000

processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : EST:*
1: gb_est1:*
2: gb_est3:*
3: gb_est4:*
4: gb_est5:*
5: gb_est6:*
6: gb_htc:*
7: gb_est2:*
8: gb_est7:*
9: gb_est8:*
10: gb_est9:*
11: gb_gss1:*
12: gb_gss2:*
13: gb_gss3:*
14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		% Query				
t o.	Score	Match	Length	DB	ID	Description
1	365	19.5	365	1	AA837817	AA837817 oe39h03.s
2	324.4	17.3	475	4	BX505056	BX505056 DKFZp686B
3	309	16.5	484	2	BG548177	BG548177 602575333
4	306.8	16.4	3143	6	HSM805999	BX537892 Homo sapi
5	305.8	16.4	648	4	CB217138	CB217138 NISC_nq12
6	305	16.3	520	8	CR556374	CR556374 DKFZp459O
7	304.8	16.3	648	14	DU636687	DU636687 Ciuffi-HI
8	304.2	16.3	679	14	AG111125	AG111125 Pan trogl
9	303.6	16.2	4640	6	CR857699	CR857699 Pongo pyg
0	303	16.2	515	9	DB062128	DB062128 DB062128
1	299.6	16.0	742	5	CF127609	CF127609 UI-HF-ET0
2	298.8	16.0	620	13	CZ450138	CZ450138 MCF730m02
3	298.8	16.0	743	3	BU629138	BU629138 UI-H-FL0-
4	298.6	16.0	720	9	CX871743	CX871743 HESC4_60_
5	297.2	15.9	537	2	BF876683	BF876683 QV0-ET014
6	296.6	15.9	583	3	BQ007997	BQ007997 UI-H-EI0-
7	295.8	15.8	588	9	DB258133	DB258133 DB258133
8	295.6	15.8	573	9	DA180845	DA180845 DA180845
9	295.2	15.8	678	14	AG118253	AG118253 Pan trogl
	20	294.8	15.8	608	7	AW970856	AW970856 EST382939
c	21	294	15.7	560	9	DA780305	DA780305 DA780305
c	22	294	15.7	8535	6	CR933630	CR933630 Homo sapi
c	23	293.6	15.7	541	4	BX508653	BX508653 DKFZp686L
c	24	293.2	15.7	489	7	BF673854	BF673854 602137512
	25	292.4	15.6	514	13	CZ456835	CZ456835 MCF739N13
	26	292.4	15.6	1084	3	BM806108	BM806108 AGENCOURT
	27	292.4	15.6	1467	6	BC039500	BC039500 Homo sapi
c	28	292.2	15.6	559	9	DB196716	DB196716 DB196716
	29	292.2	15.6	789	13	CZ461615	CZ461615 MCF750p18
	30	292.2	15.6	884	13	CZ459857	CZ459857 MCF748g07
	31	291.8	15.6	543	9	DB343878	DB343878 DB343878
	32	291.8	15.6	635	11	AQ538583	AQ538583 RPCI-11-3
	33	291.8	15.6	648	14	AG128024	AG128024 Pan trogl
c	34	291.6	15.6	496	8	CN315352	CN315352 170004237
c	35	291.4	15.6	550	9	DB102123	DB102123 DB102123
c	36	291.4	15.6	771	1	AA984258	AA984258 am83b03.s
	37	290.6	15.5	680	3	BQ181560	BQ181560 UI-H-EU0-
	38	288.8	15.4	508	13	CZ456641	CZ456641 MCF739J09
	39	288.6	15.4	517	4	CA434698	CA434698 UI-H-DI0-
c	40	288.4	15.4	579	4	BX646636	BX646636 DKFZp781K
	41	288	15.4	612	7	AW979073	AW979073 EST391183
c	42	287.8	15.4	3820	6	HSM803830	AL832522 Homo sapi
c	43	287.8	15.4	3820	6	HSM803831	AL831957 Homo sapi
	44	287.2	15.4	718	14	CR958304	CR958304 Homo sapi
c	45	287	15.3	498	2	BI551454	BI551454 603192887

ALIGNMENTS

RESULT 1
AA837817
LOCUS AA837817 365 bp mRNA linear EST 18-MAR-1998
DEFINITION oe39h03.s1 NCI_CGAP_Pr25 Homo sapiens cDNA clone IMAGE:1410965 similar to contains Alu repetitive element;contains element MER36 repetitive element ;, mRNA sequence.
ACCESSION AA837817
VERSION AA837817.1 GI:2913474
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 365)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright, Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1141 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 333.
FEATURES Location/Qualifiers
source 1. .365
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1410965"
/tissue_type="epithelium (cell line)"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_Pr25"
/note="Organ: prostate; Vector: Bluescript SK-; Site_1:

```
                EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                Oligo dT. Normal prostate epithelial cell line (HPV
                immortalized). 5' adaptor sequence: 5' GAATTCGGCACGAG 3'
                3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'
                Average insert size:  1.1 kb."
:N

ery Match             19.5%;  Score 365;  DB 1;  Length 365;
st Local Similarity   100.0%;  Pred. No. 1.9e-47;
ches  365;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

      556 GGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCG 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        1 GGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCG 60

      616 ATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       61 ATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC 120

      676 TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGA 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      121 TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGA 180

      736 TCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCT 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      181 TCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCT 240

      796 GAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAG 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      241 GAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAG 300

      856 AGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGC 915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      301 AGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGC 360

      916 TGGCC 920
          |||||
      361 TGGCC 365


;T 2
5056
:       BX505056                475 bp    mRNA    linear   EST 04-SEP-2003
IITION  DKFZp686B10192_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
        DKFZp686B10192 5', mRNA sequence.
SION    BX505056
:ON     BX505056.1  GI:32032700
)RDS    EST.
:E      Homo sapiens (human)
;ANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
(ENCE   1  (bases 1 to 475)
:HORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
        Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
        Wiemann,S.
:LE     EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
        Wellenreuther,R., et al.)
)RNAL   Unpublished (2003)
:NT     Contact: MIPS
        MIPS
        Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
        This is the 5' sequence of the clone insert
        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
        Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
        sequenced by DKFZ (German Cancer Research Center,
        Heidelberg/Germany) within the cDNA sequencing consortium of the
        German Genome Project.
        No s1 sequence available.
        This clone (DKFZp686B10192) is available at the RZPD in Berlin.
        Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
        Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             Location/Qualifiers
     source          1. .475
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686B10192"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
ORIGIN

  Query Match             17.3%;  Score 324.4;  DB 4;  Length 475;
  Best Local Similarity   82.2%;  Pred. No. 4.2e-41;
  Matches  373;  Conservative    0;  Mismatches   81;  Indels    0;  Gaps    0;

Qy        511 TTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTG 570
              ||| ||||| |||||||||||||| | |||| ||||||||||||||||||||||||| ||
Db         18 TTTTTTTTTTTTTTTTTTGAGACTGGGTCTCACTCTGTCGCCCAGGCTGGAGTGCAATG 77

Qy        571 GCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAG 630
              || |||||| |  |||  |||||||||||||| | ||||||| |||||||||||||||||
Db         78 GCGCGATCTCAGGTCACCACAACCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAG 137

Qy        631 CCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATTTTGGA 690
              |||||| |||||||||||||| |||||  | ||||||  |||||||  | || |||| ||
Db        138 CCTCCCGAGTAGCTGGGATTATAGGTGGCCACCACCATGCCCAGCTAATGTTTTTTTTGA 197

Qy        691 GACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCG 750
              ||  ||||||  | ||||| || |||||||||| || ||||  || ||| | |||||||
Db        198 GATGGAGTCTCACTCTGTCGCCGAGGCTGGAGTGCAATGGCCAGACCTCGGCTCACTGCA 257

Qy        751 ACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTA 810
              |||||  |||||||||| |||||  |||||||||||||| |||| |||||||| ||||||
Db        258 ACCTCTGCCTCCCGGGCTCAAGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTA 317

Qy        811 CAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCC 870
              |||  |||||| ||| |   || |||||||||||||||||||||| ||| | |||| ||
Db        318 CAGGCATGCACCACCAGGCCTAGATAATTTTTGTATTTTTAGTAGAAATGGGGTTTTACC 377

Qy        871 ATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCC 930
              ||||||||| ||| |||| |||| || |||||||| ||||||||  |||| |||||||
Db        378 GTGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGTTGATCCGCCTGCCTCAGCCTCCC 437

Qy        931 AAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGC 964
              |||||||||| || ||||| ||||||| | |||
Db        438 AAAGTGCTGGGCTTACAGGCCTGAGCCACCACGC 471


RESULT 3
BG548177/c
LOCUS       BG548177                484 bp    mRNA    linear   EST 04-APR-2001
DEFINITION  602575333F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703144 5',
            mRNA sequence.
ACCESSION   BG548177
VERSION     BG548177.1  GI:13546842
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 484)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
```

```
         Clone distribution: MGC clone distribution information can be
         found through the I.M.A.G.E. Consortium/LLNL at:
         http://image.llnl.gov
         Plate: LLCM1541  row: b  column: 09
         High quality sequence stop: 457.
URES             Location/Qualifiers
 source          1. .484
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:4703144"
                 /lab_host="DH10B (T1 phage-resistant)"
                 /clone_lib="NIH_MGC_77"
                 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                 SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                 3' adaptors were used in cloning as follows: 5' adaptor
                 sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                 C, or G and N = A, C,  G, or T).  Average insert size 1.9
                 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                 by PCR. This library was enriched for full-length clones
                 and was constructed by Clontech  Laboratories (Palo Alto,
                 CA). Note: this is a NIH_MGC Library."
:N

ery Match              16.5%;  Score 309;  DB 2;  Length 484;
st Local Similarity    80.4%;  Pred. No. 1.1e-38;
tches  377;  Conservative    0;  Mismatches   85;  Indels    7;  Gaps    1;

       502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
           | || || | || ||| |  | ||||||||||||||||||  |||| || ||||||.|||
       469 TCTTATTCTCTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 410

       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
           | ||||||||  ||||||  || ||| ||||||||||||||||| | |||||| |||||||
       409 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 350

       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
           ||||||||||||||  |||||||||| |||||||| | |||| || ||| ||
       349 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 290

       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
            || |||| |||| |||| ||| ||  | |||| |||||||||||| || |||||||
       289 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 230

       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
           ||||||| ||||||| |||||||||| || |||||||| |||||||||||||||| ||||
       229 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 170

       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
           ||||||||  |||||||||  |  | || |||||| | ||||||||| ||||||||||
       169 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 110

       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
           ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
       109 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 50

       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
           |  | ||||||  ||| ||||||||||||||| |||| |||||||| ||||
        49 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCG 1


'LT 4
'05999
S        HSM805999               3143 bp    mRNA    linear   HTC 22-SEP-2004
NITION   Homo sapiens mRNA; cDNA DKFZp313M1136 (from clone DKFZp313M1136).
SSION    BX537892
:ON      BX537892.1  GI:31873905
'ORDS    HTC.
CE       Homo sapiens (human)
GANISM   Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
         Hominidae; Homo.
REFERENCE   1  (bases 1 to 3143)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  CONSRTM   The German cDNA Consortium
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
            Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the
            Heinrich-Heine-University, Duesseldorf/Germany) within the cDNA
            sequencing consortium of the German Genome Project. This clone
            (DKFZp313M1136) is available at the RZPD Deutsches
            Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
            Please contact RZPD for ordering:
            http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp313M1136
            Further information about the clone and the sequencing project is
            available at http://mips.gsf.de/projects/cdna/.
FEATURES             Location/Qualifiers
     source          1. .3143
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="RZPD:DKFZp313M1136"
                     /db_xref="taxon:9606"
                     /clone="DKFZp313M1136"
                     /tissue_type="heart"
                     /clone_lib="313 (synonym: hlcc2). Vector pTriplEx2; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
                     /note="chimeric"
ORIGIN

  Query Match             16.4%;  Score 306.8;  DB 6;  Length 3143;
  Best Local Similarity   80.8%;  Pred. No. 1.3e-38;
  Matches  375;  Conservative    0;  Mismatches   77;  Indels   12;  Gaps    1;

Qy        518 TTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGAT 577
              || ||||||||||||||||||||||||| |||||| |||||||||||||||||||||||||
Db       1828 TTTTTTTTTTTTGAGACAGAGTCTTACTCTGTTGCCCAGGCTGGAGTGCAGTGGCATGAT 1887

Qy        578 CTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCA 637
              ||| || |||||||  |||  |||||| | |||||| |||||||||| ||||||||||
Db       1888 CTCGGCTCACTGCAGACTCCACCTCCCAGGTTCAAGTGATTCTCCTGTCTCAGCCTCCCG 1947

Qy        638 AGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------------TTTTTTTATT 685
              ||||||||||||||||||     |  |||||   |||||           ||||||| ||
Db       1948 AGTAGCTGGGATTACAGGCATGCAGCACCATGTCCAGCTAATTTTCTGTATTTTTTTTTT 2007

Qy        686 TTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCA 745
              || ||||| |||||| || ||||  |||||||||||||| |||||||| |||||||| |||
Db       2008 TTTGAGACGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGTGATCTCAGCTCA 2067

Qy        746 CTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAG 805
              |||| ||||||  ||||| ||||||||||||||||||||||||||| |||||||||||| |
Db       2068 CTGCAACCTCCGTCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCTGAGTAGCTGG 2127

Qy        806 GATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTT 865
              || |||||    |  |  |||||    |||||||||| |||||||||||||| | | ||
Db       2128 GACTACAGGCACGTGCTACCACGCCTGGCTAATTTTTTTATTTTTAGTAGAGACGGGGTT 2187

Qy        866 TTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGC 925
              || |||| |||| || | | |||| ||| || |||||||||||||| |  | ||  ||
Db       2188 TTACCATATTGGTCAGGTTGGTCTTGAACTCCTGACCTCAGGTGATCCACCCGTCTCAGC 2247

Qy        926 CTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGC 969
              |||||||||||||||||| ||||||||||||  || || | |
Db       2248 CTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCTGAC 2291


RESULT 5
CB217138/c
LOCUS       CB217138                 648 bp    mRNA    linear   EST 06-FEB-2003
```

```
ITION  NISC_nq12g07.y1 NICHD_HS_Ut2 Homo sapiens cDNA clone IMAGE:5938765
       5', mRNA sequence.
SION   CB217138
ON     CB217138.1  GI:28265330
RDS    EST.
E      Homo sapiens (human)
ANISM  Homo sapiens
       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
       Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
       Hominidae; Homo.
ENCE   1  (bases 1 to 648)
HORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
LE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
       Tumor Gene Index
RNAL   Unpublished (1997)
NT     Contact: Robert Strausberg, Ph.D.
       Email: cgapbs-r@mail.nih.gov
       cDNA Library Preparation:
        cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
        DNA Sequencing by: National Institutes of Health Intramural
       Sequencing Center (NISC)
        Clone distribution: NCI-CGAP clone distribution information can be
       found through the I.M.A.G.E. Consortium/LLNL at:
       info@image.llnl.gov
       Plate: LLAM13167  row: N  column: 14
       Seq primer: M13RP1 reverse primer (ABI).
RES            Location/Qualifiers
 source        1. .648
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5938765"
               /sex="female"
               /tissue_type="normal endometrium, mid-secretory phase,
               cycle day 23"
               /lab_host="DH10B (T1-resistant)"
               /clone_lib="NICHD_HS_Ut2"
               /note="Organ: uterus; Vector: pCMV-SPORT6.1.ccdb (ResGen,
               Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
               Cloned unidirectionally from microquantity amounts of mRNA
               from normal endometrial tissue (mid-secretory phase, cycle
               day 23). Average insert size 1.6 kb. Library constructed
               by ResGen (Invitrogen Corporation)."
N

ry Match            16.4%;  Score 305.8;  DB 4;  Length 648;
t Local Similarity  82.1%;  Pred. No. 3.1e-38;
ches  380;  Conservative    0;  Mismatches   72;  Indels   11;  Gaps    2;

Qy        509 TTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAG 568
              ||||  || || |||||||| |||||||||| |     ||||||||||||||||||||
Db        462 TTTTCTTTCTTTTTTTTTTGGAGACAGAGTTTGCTCTTGTCGCCCAGGCTGGAGTGCAA 403

Qy        569 TGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTC 628
              ||||| | |||| || |||||||||||||||||||||| |||||| ||||||||||||||
Db        402 TGGCACGCTCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCTC 343

Qy        629 AGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----------TTT 678
              |||||||| |||||||||| |||||||||    ||||||||||||||          |||
Db        342 AGCCTCCCGAGTAGCTGGGGTTACAGGTGTGTGCCACCACACCCAGCTAATTTTTCTTTT 283

Qy        679 TTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCT 738
              ||| ||||  |||  ||||||   |  |||| | ||||||||||||||||||| ||||||
Db        282 CTTTTTTTTTAAGAGGGAGTCTCACTGTGTCGC-CAGGCTGGAGTACAGTGGCGTGATCT 224

Qy        739 CAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAG 798
              | ||||||| |||||    |||| ||||||||||||||||||||||||||||||| |||
Db        223 TGGCTCACTGCAACCTCTGACTCCTGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCCGAG 164

Qy        799 TAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGA 858
              ||||| |||||||||  | |  || ||||   |||||||||||||||||||||||||||
Db        163 TAGCTGGGATTACAGGTGCGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGA 104

Qy        859 TGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGG 918
              || | |||  |||||| ||||| ||| |||| |||| || ||||||||||||||| ||
Db        103 TGGGGTTTCACCATGTCGGCCAAGCTGGTCTCGAACTCCTGACCTCAGGTGATCTGCCCA 44

Qy        919 CCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCG 961
              || ||||||||||||||||||||||| | ||||||||||| ||
Db         43 CCCTGGCCTCCCAAAGTGCTGGGATTACGGGCGTGAGCCACCG 1


RESULT 6
CR556374/c
LOCUS       CR556374                520 bp    mRNA    linear   EST 12-JUL-2004
DEFINITION  DKFZp459O0431_r1 459 (synonym: pcor1) Pongo pygmaeus cDNA clone
            DKFZp459O0431 5', mRNA sequence.
ACCESSION   CR556374
VERSION     CR556374.1  GI:50249976
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pongo.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Koehrer,K., Beyer,A., Mewes,H.W., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert Clone from S. Wiemann,
            Molecular Genome Analysis, German Cancer Research Center (DKFZ);
            Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical
            Research Center at the Heinrich-Heine-University,
            Duesseldorf/Germany) within the cDNA sequencing consortium of the
            German Genome Project. This clone (DKFZp459O0431) is available at
            the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
            Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
            clone@rzpd.de Further information about the clone and the
            sequencing project is available at
            http://mips.gsf.de/projects/cdna/.
FEATURES            Location/Qualifiers
     source         1. .520
                    /organism="Pongo pygmaeus"
                    /mol_type="mRNA"
                    /db_xref="taxon:9600"
                    /clone="DKFZp459O0431"
                    /tissue_type="cortex"
                    /dev_stage="adult"
                    /lab_host="DH10B"
                    /clone_lib="459 (synonym: pcor1)"
                    /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

  Query Match            16.3%;  Score 305;  DB 8;  Length 520;
  Best Local Similarity  80.0%;  Pred. No. 4.5e-38;
  Matches  385;  Conservative    0;  Mismatches   87;  Indels    9;  Gaps    2;

Qy        498 CTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGG 557
              ||   | | | | | ||| | |  || |||||| |||| |||| |||||||| ||||||
Db        491 CTATCTATCTATCTATCTATCTAATTATTTTTGNGACACGGTCTCGCTCTGTCACCCAGG 432

Qy        558 CTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGAT 617
              ||||||||||||||  ||||||| || ||||||| |||| |||||||||  ||||||||
Db        431 CTGGAGTGCAGTGGTGTGATCTCGGCTCACTGCAGCCTCCGCCTCCCGGGCTCAAGCGAN 372

Qy        618 TCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-- 675
              ||||| |||||||||||||||||||||||| ||||||  | | ||||||  ||  ||
Db        371 TCTCCCGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCACCCACCACCATGCCTGGCTA 312

Qy        676 -----TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGG 730
                  ||||||| ||   |||||||||||||| ||||| |||||||||||| | ||||
Db        311 ATTTTTTTTTTTTTTCCTGAGACAGAGTCTTGCTCTGTCCCCCAGGCTGGAGTGCGGTGG 252
```

```
      731 CATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGT 790
          ||  ||||||| |||||| |||||| |||||| ||||||||| ||||||||||||||||
      251 CACAATCTCAGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGC 192

      791 CTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTT 850
          |||| ||||||||||||||||||   ||||||   | ||   ||||||||||||||||||
      191 CTCCCGAGTAGCTAGGATTACAGGCATGCACCATGAAGTCTGGCTAATTTTTGTATTTTT 132

      851 AGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGA 910
          |||||||||| |  ||  |||||||||||| ||| |||| |||| || ||||||  ||||
      131 AGTAGAGATGGGGCTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTC--GTGA 74

      911 TCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          ||| ||  |||||||||||||||| ||||||| | |||||  ||| || | ||||  |||
       73 TCCACTCACCTTGGCCTCCCAAAGCGCTGGGACTACAGGCTGGAGTCACCACGCCCAGCC 14

      971 T 971
          |
       13 T 13
```

```
LT 7
6687/c
S       DU636687                 648 bp    DNA     linear   GSS 27-OCT-2005
NITION  Ciuffi-HIV-HOS-sip75-1F10.fwdQ=52 Human Integration Site
        Library-Ciuffi-HIV-HOS-siL Homo sapiens genomic, genomic survey
        sequence.
SSION   DU636687
ION     DU636687.1  GI:78199057
ORDS    GSS.
CE      Homo sapiens (human)
GANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
ENCE    1  (bases 1 to 648)
THORS   Ciuffi,A., Llano,M., Poeschla,E., Hoffmann,C., Leipzig,J.,
        Shinn,P., Ecker,J.R. and Bushman,F.D.
TLE     A role for LEDGF/p75 in targeting HIV DNA integration
URNAL   Nat. Med. (2005) In press
NT      Contact: Bushman FD
        Department of Microbiology
        University of Pennsylvania School of Medicine
        402C Johnson Pavilion, 3610 Hamilton Walk, Philadelphia, PA
        19104-6076, USA
        Tel: 215 573 8732
        Fax: 215 573 4856
        Email: bushman@mail.med.upenn.edu
        The hg17 (May 2004) freeze of the human genome was used.
        Class: shotgun.
URES            Location/Qualifiers
 source         1. .648
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /cell_type="HOS-siL (HOS transduced with a retroviral
                construct containing a shRNA targeting LEDGF/p75)"
                /clone_lib="Human Integration Site
                Library-Ciuffi-HIV-HOS-siL"
                /note="Sequences cloned using TOPO vectors."
N
```

```
ery Match             16.3%;  Score 304.8;  DB 14;  Length 648;
st Local Similarity   81.4%;  Pred. No. 4.5e-38;
tches  390;  Conservative    0;  Mismatches   83;  Indels    6;  Gaps    3;

      496 TGCTGCTTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCA 555
          | |  |||| |||||||| ||||| ||||||||||||||||||||||||||||||||||
      588 TACCACTTTCTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCA 529

      556 GGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCG 615
          |||||||||||||||| ||||||     |||||||| |||  | |||||| ||||  |
```

```
Db        528 GGCTGGAGTGCAGTGGCGTGATCTTGCNTCACTGCAAGCTCGCCTTCCCGGGTTCACACC 469

Qy        616 ATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC 675
              |||||||||||||||||||||| |||||||||| ||||||||  | ||||||||||||||
Db        468 ATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGCCCACCACCACACCCAGC 409

Qy        676 T--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCAT 733
              |  ||||||  |||||      |||||  || ||||||||||||||||||| ||||     |
Db        408 TAATTTTTTTTTTTTGAGACGGGAGTCACGCTCTGTCACCCAGGCTGGAGTGCAGTCATGT 349

Qy        734 GATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTC 793
              |||||||| ||||||| | |||| |||||||||||||| || |||||||||||||||| |||
Db        348 GATCTCAGCTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTC 289

Qy        794 CTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--GTATTTTTA 851
              | |||||||| ||| |||||  |  | || |||||  | ||||||||||     || |||
Db        288 CCGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTTGTATTTCTTA 229

Qy        852 GTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGAT 911
              ||||||| | | |||  ||||||| |||| | | |||| || | || ||||||  |||||
Db        228 GTAGAGACGAGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTC--GTGAT 171

Qy        912 CCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
              || |  |||||||||||||||||||||||||||||  |||||||||||| | |||| ||||
Db        170 CCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTATAGGCGTGAGCCACCACGCCCGGCC 112
```

```
RESULT 8
AG111125
LOCUS       AG111125                 679 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-117A23.F, genomic survey sequence.
ACCESSION   AG111125
VERSION     AG111125.1  GI:16731644
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 679)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
             PRIMERS
                Sequencing: -21M13
             LIBRARY
                Vector    : pKS145
                R.Site 1  : SacI
                R.Site 2  : SacI.
FEATURES            Location/Qualifiers
     source         1. .679
                    /organism="Pan troglodytes"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9598"
                    /clone="PTB-117A23.F"
                    /sex="male"
                    /cell_type="lymphoblast"
                    /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
```

```
:ry Match            16.3%;  Score 304.2;  DB 14;  Length 679;
:t Local Similarity   78.9%;  Pred. No. 5.5e-38;
:ches  375;  Conservative    0;  Mismatches   98;  Indels    2;  Gaps    1;

      499 TGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGC 558
          | | | | | | | | | | ||||||||||||||||||||  |||||||  ||||||||
      179 TACATATATATATATATATATTTTTTTTTTTGAGACAGAGTCCCGCTCTGTGGCCCAGGC 238

      559 TGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATT 618
          |||||||||||||  |||||| || ||||||||  |||  ||||||||  |||| || |||
      239 TGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCACCTCCCGGGTTCACGCCATT 298

      619 CTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC--T 676
          ||||||||||||||||  ||||||||||| ||||||  | |||||||||  | | ||
      299 CTCCTGCCTCAGCCTCCGGAGTAGCTGGGACTACAGGCACCCGCCACCACGCACGGCCTA 358

      677 TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          | | |  |||| |||||||| |||  | |||||  |||||  |||||||  ||||||||||||
      359 TATATATTTTTTGAGACAGAATCTCACTCTGTCGCCCAGTCTGGAGTGCAGTGGCATGAT 418

      737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
          ||  | |||||||  ||||||  ||||||||||||||||  ||||||||||||||| |||| |
      419 CTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 478

      797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGA 856
          | |||||  |||  ||||  | ||  ||  |||  |  |||||||||||  ||  |||||||
      479 AATAGCTGGGACAACAGGTGCGCGCCACCATGACCCGCTAATTTTTGTTGTTGTAGTAGA 538

      857 GATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCT 916
          ||   | |||  |||||||||||| ||| |||| |||| || ||||||||||||| ||
      539 GACAGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTGCC 598

      917 GGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
           |||| |||| |||||||||||||||||| ||||| ||| ||  |  | |  || |
      599 TGCCTCGGCCCCCCAAAGTGCTGGGATTACAGGCCTGATCCTCCCTGGCCCGCTT 653


;T 9
'699
;       CR857699                4640 bp    RNA     linear   HTC 11-JAN-2005
)ITION  Pongo pygmaeus mRNA; cDNA DKFZp459J241 (from clone DKFZp459J241).
;SION   CR857699
:ON     CR857699.1  GI:55726396
)RDS    HTC.
:E      Pongo pygmaeus (orangutan)
;ANISM  Pongo pygmaeus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Pongo.
:ENCE   1  (bases 1 to 4640)
'HORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
        Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
        Wiemann,S.
:SRTM   The German cDNA Consortium
'LE     Direct Submission
:RNAL   Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
        Neuherberg, GERMANY
;NT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
        Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
        sequenced by DKFZ (German Cancer Research Center,
        Heidelberg/Germany) within the cDNA sequencing consortium of the
        German Genome Project. This clone (DKFZp459J241) is available at
        the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
        Berlin, Germany. Please contact RZPD for ordering:
        http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp459J241
        Further information about the clone and the sequencing project is
        available at http://mips.gsf.de/projects/cdna/.
)RES            Location/Qualifiers
 source         1. .4640
                /organism="Pongo pygmaeus"
                /mol_type="pre-RNA"
                /db_xref="taxon:9600"
                /clone="DKFZp459J241"
                /tissue_type="cortex"
                /clone_lib="459 (synonym: pcor1). Vector pSport1_Sfi; host
                DH10B; sites SfilA + SfilB"
                /dev_stage="adult"
                /note="hypothetical protein (Homo sapiens), C-terminus
                truncated, not fully spliced"
      gene      <128. .>610
                /gene="DKFZp459J241"
      CDS       <128. .>610
                /gene="DKFZp459J241"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAH89968.1"
                /db_xref="GI:55726397"
                /translation="SQFKRQRINPLPGGRNFSGTASTSLLGPPPGLLTPPVATELSQN
                ARHLQGGEKQRVFTGIVTSLHDCFGVVDEEVFFQLSVVKGRLPQLGEKVLVKAAYNPG
                QAVPWNAVKVQTLSNQPLLKSPAPPLLHVAALGQKQGILGAQPQLIFRPHRIPPLFPQ
                K"
ORIGIN

  Query Match            16.2%;  Score 303.6;  DB 6;  Length 4640;
  Best Local Similarity  80.6%;  Pred. No. 3.6e-38;
  Matches  379;  Conservative    0;  Mismatches   89;  Indels    2;  Gaps    2;

Qy        504 TTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAG 563
              ||||| |||| ||| | | |||||||  ||  ||||||  |||  ||| ||||||||||||
Db        754 TTTTTGTTTTGTTTCTTTCTTTTTTTTTGATGGAGTCTCGCTGTGTTGCCCAGGCTGGAG 813

Qy        564 TGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCT 623
              ||||||| || |||||| || ||||||||||||||||||| || |||||||  ||| |  |
Db        814 TGCAGTGACACGATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTATTTT 873

Qy        624 GCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTA 683
              |||||||  ||||   |||||||||||||||| ||  |||||||  |||  |||  ||||
Db        874 GCCTCAGTTTCCCCCTTAGCTGGGATTACAGGCGCCTGCCACCATGCCCGGCTAATTTT- 932

Qy        684 TTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTT 743
              ||||  |||||||||||| || ||||||||||||||||||||||||||||  |||||| | |
Db        933 TTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTGGAGTACAGTGGCGCGATCTCGGCT 992

Qy        744 CACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCT 803
              || ||| | |||| ||||| |||||||  | ||||||||||||||| ||| |||||||||
Db        993 CATTGCAAGCTCCGCCTCCTGGGTTCACACCATTCTCCTGCCTCAGCCTCTTGAGTAGCT 1052

Qy        804 AGGATTACAGAAGTGCACCTCCACGTTCAGCTAA-TTTTTGTATTTTTAGTAGAGATGCG 862
              ||| |||||  |  |||| |||||  | |||||  ||||||||||||||||||||| | |
Db       1053 GGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGG 1112

Qy        863 CTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTT 922
              |||  ||||||| |||| |||| |||| || ||||||| |||||||     || ||
Db       1113 GTTTCACCATGTTCGCCAGGCTGGTCTTGAACTCCTGACCTCAAGTGATCCATCTGCATT 1172

Qy        923 GGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTG 972
              | |||| |||||||||||||| |||| |||||| | ||  || |||| |
Db       1173 GTCCTCTCAAAGTGCTGGGATTACAGGTGTGAGCTACCGTACCCGGCCGG 1222


RESULT 10
DB062128/c
LOCUS       DB062128                515 bp    mRNA    linear   EST 10-DEC-2005
DEFINITION  DB062128 TESTI4 Homo sapiens cDNA clone TESTI4002179 5', mRNA
            sequence.
ACCESSION   DB062128
VERSION     DB062128.1  GI:83451015
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Kimura,K., Wakamatsu,A., Suzuki,Y., Ota,T., Nishikawa,T.,
            Yamashita,R., Yamamoto,J., Sekine,M., Tsuritani,K., Wakaguri,H.,
```

```
        Ishii,S., Sugiyama,T., Saito,K., Isono,Y., Irie,R., Kushida,N.,
        Yoneyama,T., Otsuka,R., Kanda,K., Yokoi,T., Kondo,H., Wagatsuma,M.,
        Murakawa,K., Ishida,S., Ishibashi,T., Takahashi-Fujii,A.,
        Tanase,T., Nagai,K., Kikuchi,H., Nakai,K., Isogai,T. and Sugano,S.
TLE     Diversification of Transcriptional Modulation: Large-scale
        Identification and Characterization of Putative Alternative
        Promoters of Human Genes
URNAL   Genome Res. 16 (1), 55-65 (2006)
UBMED   16344560
NT      Contact: Takao Isogai
        FLJ Project (HRI Team)
        Helix Research Institute
        2-6-7 Kazusa-Kamatari, Kisarazu, Chiba, 292-0818, Japan
        Tel: 81-438-52-3975
        Fax: 81-438-52-3986
        Email: flj-cdna@nifty.com
        NEDO human cDNA project (New Energy and Industrial Technology
        Developmental Organization, Japan); cDNA library construction:
        Helix Research Institute (HRI); 5'-end one pass sequencing: HRI,
        Research Association for Biotechnology (RAB) and Biotechnology
        Center, National Institute of Technology and Evaluation; 3'-end one
        pass sequencing: RAB.
URES            Location/Qualifiers
 source         1. .515
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="TESTI4002179"
                /tissue_type="testis"
                /clone_lib="TESTI4"
                /note="Vector: pME18SFL3"
N

ery Match            16.2%;  Score 303;  DB 9;  Length 515;
st Local Similarity  80.8%;  Pred. No. 9.3e-38;
tches  392;  Conservative    0;  Mismatches   85;  Indels    8;  Gaps    3;

      508 TTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTC-TGTCGCCCAGGCTGGAGTGC 566
          |||   |||||| ||||||||||||||| |||| |||||| ||| ||||||||||||| |
      493 TTTCACCTTTTTTTTTTTTTTTGAGACGGAGTTTGCTCTTGTTGCCCAGGCTGGAGTCC 434

      567 AGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCC 626
          |||| || |||||  || || ||||||||  |||||| || ||||||||||||||||||||
      433 AGTGACACGATCTTGGCTCAATGCAACCTCCACCTCCAGGGTTCAAGCGATTCTCCTGCC 374

      627 TCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------TTTTT 680
          |||| ||||  |||||||| |||||||||  |  ||||||||||||||      |||||
      373 TCAGTCTCCTGAGTAGCTGAGATTACAGGCACGTGCCACCACACCCAGCTAATTTTTTTT 314

      681 TTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCA 740
          || | || |||||||||||| || ||||| |||||||||||| |||||||  |||||||
      313 TTTTCTTTGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCA 254

      741 GTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTA 800
          | ||||||| |||||| ||||||||| |||||| |||||||||||||||| | |||||||
      253 GCTCACTGCAACCTCCGCCTCCCGGGCTCAAGCGATTCTCCTGCCTCAGTGTTCTGAGTA 194

      801 GCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATG 860
          ||| ||| |||||| ||    | ||| |  |||||||||||| |||||||||||||| |
      193 GCTGGGACTACAGACGTATGTCACCATGCCCAGCTAATTTTTGCATTTTTAGTAGAGACG 134

      861 CGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCC 920
          | |||  |||||||||||| |||  ||   ||| || |||||||||||||| |  |||
      133 GGGTTTCACCATGTTGGCCAGGCTGATCGCAAACTCCTGACCTCAGGTGATCCACCCGCC 74

      921 TTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTC 980
          ||||||| |||||||| ||||||| ||||  |||||||  |||||  |||||      | |
       73 TTGGCCT-CCAAAGTGTTGGGATTACAGGTATGAGCCAAAGCGCCCGGCCTCTCTCTCAC 15

      981 CTTTA 985
          ||||
       14 TTTTA 10
```

```
RESULT 11
CF127609
LOCUS       CF127609                742 bp    mRNA    linear   EST 05-AUG-2003
DEFINITION  UI-HF-ET0-awi-i-03-0-UI.r1 NIH_MGC_214 Homo sapiens cDNA clone
            IMAGE:30553058 5', mRNA sequence.
ACCESSION   CF127609
VERSION     CF127609.1  GI:33206022
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 742)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  PUBMED    8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Mary Hendrix
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/humanfl.html
             The following repetitive elements were found in this cDNA
            sequence: 69-280, >ALU (matched compliment) 241-528, >ALU (matched
            compliment) 531-635, >MER63B#DNA/MER1_type? 531-722,
            >MER63A#DNA/MER1_type?
            Seq primer: pYX-5.
FEATURES            Location/Qualifiers
     source         1. .742
                    /organism="Homo sapiens"
                    /mol_type="mRNA"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:30553058"
                    /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
                    /lab_host="DH10B (T1 phage resistant)"
                    /clone_lib="NIH_MGC_214"
                    /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
                    Site_2: Not I; The library was constructed according
                    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                    1996. Denatured RNA was size fractionated on a 1% agarose
                    gel. First strand cDNA synthesis was primed with oligo-dT
                    primer containing a Not I site. Double strand cDNA was
                    size selected according to mRNA size fraction, ligated
                    with EcoR I adaptor, digested with Not I and then cloned
                    directionally into pYX-Asc vector. The library tag
                    sequence located between the Not I site and the polyA tail
                    is GATAAGGCCA. Tissue was provided by Mary Hendrix."
ORIGIN

  Query Match            16.0%;  Score 299.6;  DB 5;  Length 742;
  Best Local Similarity  78.2%;  Pred. No. 2.8e-37;
  Matches  391;  Conservative    0;  Mismatches   94;  Indels   15;  Gaps    2;

Qy        499 TGCTTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGC 558
              || | | ||      ||||| |||||||||||||| |||||| ||||||| |||||||
Db         46 TGGTATATTGGGGAATTTTTTTTTTTTTTTTGAGACGGAGTCTGGCTCTGTCACCCAGGC 105

Qy        559 TGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATT 618
              | ||||||| ||||| |||||| || |||||||| |||||| ||| || || | || |||
Db        106 TAGAGTGCAATGGCACGATCTCGGCTCACTGCAAGCTCTGCGTCCTGGGTTGACGCCATT 165

Qy        619 CTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC--- 675
```

```
          |||||||||||||||||| |||||||||| |||||| || ||||||||||||||||
      166 CTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGAGCTCGCCACCACACCCAGCTAA 225

      676 ----------TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
                    ||||||| | || ||||| |||| | || ||||| |||||||||||||| |
      226 TTTTTTTTTTTTTTTTTTTATTTGAGACGGAGTTTCGCTCTGTCGCCCAGGCTGGAGTGC 285

      726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
          ||||||  |||||||   ||||||| | |||| |||||||||||||| || ||||||||||
      286 AGTGGCGCGATCTCGACTCACTGCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCC 345

      786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTA 845
          |||| |||| | |||||| ||| |||||    || || ||| |  | ||||||||||||||
      346 TCAGCCTCCCGTGTAGCTGGGACTACAGGCACGCGCCACCATGCCCGGCTAATTTTTGTA 405

      846 TTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCA 905
          |||||||||||||||| | |||  || |||| |||| | | |||| || | || |||||
      406 TTTTTAGTAGAGATGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTGACCT-- 463

      906 GGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCC 965
          |||||||| |   |||||||||||||||||||||||||| |||||||||||||| ||| ||
      464 GGTGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACC 523

      966 AGGCCTGAGCTACTCCTTTA 985
            ||| | |  | |  | ||
      524 CAGCCAGGGGGAATGTTCTA 543
```

```
LT 12
:138
:        CZ450138                620 bp    DNA     linear   GSS 20-OCT-2005
:ITION   MCF730m02TF Human MCF7 breast cancer cell line library (MCF7_1)
         Homo sapiens genomic clone MCF7_30m02, genomic survey sequence.
:SION    CZ450138
:ON      CZ450138.1  GI:77926400
:RDS     GSS.
:E       Homo sapiens (human)
:ANISM   Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
         Hominidae; Homo.
:ENCE    1  (bases 1 to 620)
:HORS    Volik,S.V., Raphael,B.J., Huang,G.-Q., Murnane,J., Brebner,J.H.,
         Bajsarowicz,K., Paris,P., Tao,Q., Kowbel,D., Lapuk,A.V., Kuo,W.-L.,
         Shagin,D.A., Shagina,I.A., Magrane,G., Gray,J.W., Jan,F.-C., de
         Jong,P., Pevzner,P. and Collins,C.
:LE      Decoding the genomic architecture and high throughput detection of
         fusion transcripts in breast cancer cell lines: implications for a
         tumor genome project
:RNAL    Unpublished (2005)
:NT      Contact: Volik SV
         Colin Collins' lab
         UCSF Comprehensive Cancer Center
         UCSF Box 0808, San Francisco, CA 94143-0808, USA
         Tel: 415 502 7066
         Fax: 415 502 5665
         Email: svolik@cc.ucsf.edu
         This clone is available from Amplicon Express
         http://www.genomex.com
         Class: BAC ends.
:RES             Location/Qualifiers
 source          1. .620
                 /organism="Homo sapiens"
                 /mol_type="genomic DNA"
                 /db_xref="taxon:9606"
                 /clone="MCF7_30m02"
                 /sex="female"
                 /clone_lib="Human MCF7 breast cancer cell line library
                 (MCF7_1)"
                 /note="Vector: pECBAC1; Site_1: HindIII; This library was
                 constructed from MCF7 breast cancer cell line by Amplicon
                 Express (http://www.genomex.com) using their standard
                 procedure."
```

```
ORIGIN

  Query Match           16.0%;  Score 298.8;  DB 13;  Length 620;
  Best Local Similarity   83.8%;  Pred. No. 4e-37;
  Matches  362;  Conservative    0;  Mismatches   67;  Indels    3;  Gaps    2;

Qy        536 GAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCT 595
              |||||| |||||||| |||||| |||||||||||||    ||||| || ||||||||||
Db        192 GAGTCTCGCTCTGTCACCCAGGTTGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCT 251

Qy        596 CTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGG 655
              | ||||||| | |||||| |||||||||||||||||||||| |||||||||||||| |||
Db        252 CCGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGGATTATAGG 311

Qy        656 TGCACGCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAG 715
                | | |||||| |||||||| |||| ||||  |||| ||||||||| ||||| |||||
Db        312 CACCCACCACCAGACCCAGCTAATTTTCTTTT--AGACGGAGTCTTGCTCTGTCGCCCAG 369

Qy        716 GCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAA 775
              |||||||| |||||||| |||||| | |||||| |||||||||||| ||||||||||||
Db        370 GCTGGAGTGCAGTGGCACGATCTCGGCTCACTGCTACCTCCACCTCCTGGGTTCAAGCAA 429

Qy        776 TTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCT 835
              |||||||| ||||| |||||||||||| | ||||||    ||  || |||||  |||||
Db        430 TTCTCCTGTCTCAGCCTCCTGAGTAGCTGGAATTACATGCATGTGCCACCACGCCCAGCT 489

Qy        836 AATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACC 895
              |||||||||||||||||||||||||| |||||  |||||  ||||||  ||  |||| ||||
Db        490 AATTTTTGTATTTTTAGTAGAGATGGGCTTTCACCATGCTGGCCAGACTGGTCTCGAACT 549

Qy        896 CCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAG 955
              |  |||||| |||||| ||    ||| |||||||||||||||| |||| ||| | ||||
Db        550 CTTGACCTCA-GTGATCTGCCCACCTCGGCCTCCCAAAGTGCTGAGATTACAGACATGAG 608

Qy        956 CCATCGCGCCAG 967
              |||  |||||  |
Db        609 CCACAGCGCCTG 620


RESULT 13
BU629138
LOCUS       BU629138                743 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  UI-H-FL0-bdh-l-10-0-UI.s1 NCI_CGAP_FL0 Homo sapiens cDNA clone
            UI-H-FL0-bdh-l-10-0-UI 3', mRNA sequence.
ACCESSION   BU629138
VERSION     BU629138.1  GI:23295352
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
             The following repetitive elements were found in this cDNA
            sequence: 11-338, >ALU (matched compliment) 185-474, >ALU (matched
            compliment) 390-613, >ALU (matched compliment)
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES            Location/Qualifiers
     source         1. .743
                    /organism="Homo sapiens"
```

```
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-H-FL0-bdh-l-10-0-UI"
                /tissue_type="Cell lines"
                /dev_stage="Adult"
                /lab_host="DH10B (Life Technologies)"
                /clone_lib="NCI_CGAP_FL0"
                /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
                (Pharmacia) with a modified polylinker; Site_1: EcoR I;
                Site_2: Not I; NCI_CGAP_FL0 is a cDNA library derived from
                a pool of mRNA obtained from 4 cell lines from grade III
                chondrosarcoma tissues. The library was constructed
                according to Bonaldo, Lennon and Soares, Genome Research,
                6:791-806, 1996. First strand cDNA synthesis was primed
                with an oligo-dT primer containing a Not I site. Double
                stranded cDNA was ligated to an EcoR I adaptor, digested
                with Not I, and cloned directionally into pT7T3-Pac
                vector. The oligonucleotide used to prime the synthesis of
                first-strand cDNA contains a library tag sequence that is
                located between the Not I site and the (dT)18 tail. The
                sequence tag for this library is GAGGTCGGTG. The cell line
                was provided by Dr James Martin from University of Iowa.
                TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
                TAG_LIB=UI-H-FL0
                TAG_SEQ=GAGGTCGGTG"
:N

ery Match            16.0%;  Score 298.8;  DB 3;  Length 743;
st Local Similarity  76.1%;  Pred. No. 3.8e-37;
tches  388;  Conservative    0;  Mismatches  107;  Indels   15;  Gaps    1;

      512 TTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGG 571
          || ||||| |||||||||||||||||||  |||||| ||||||||||||||||||||
        1 TTTTTTTTTTTTTTTTTTGAGACAGAGTCTCACTCTGTTGCCCAGGCTGGAGTGCAGTGA 60

      572 CATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGC 631
          |   |||||||| |||||||| ||| |||||| || |||| || |||||||||||| ||
       61 CGCAATCTCTGCTCACTGCAAGCTCCGCCTCCTGGGTTCATGCCATTCTCCTGCCTCTGC 120

      632 CTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---------------T 676
          ||||| |||||||||||| |||||| || | |||| ||||   ||               |
      121 CTCCCGAGTAGCTGGGACTACAGGCGCCCACCACTACACTAGGCAAATTTTTTGCATTTT 180

      677 TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          ||  || |||||||  |||||||| || || || |||||||||||| |  ||||   ||
      181 TTAATTTTTTTGGATGCAGAGTCTCGCTCTATCCCCCAGGCTGGAGTGCGCTGGCGCAAT 240

      737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
          ||  | ||||||| |||||  |||||| |||||||  |||||||||||||| ||||||
      241 CTTGGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTG 300

      797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGA 856
          |||| || |||||||||  |    |  |||||  ||||||||||||||||||||||||||
      301 AGTATCTGGGATTACAGGTGCCTGCAACCACGCCCAGCTAATTTTTGTATTTTTAGTAGA 360

      857 GATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCT 916
          |||| | ||| |||||||  || || ||| |||| |||| || |||||| ||||||||
      361 GATGGGGTTTCGCCATGCCGGTCAGGCTGGTCTTGAACTCCTGACCTCGGGTGATCCTTC 420

      917 GGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCT 976
          |||| |||||||||||||||||||||| |||||||||  || || ||| ||||      |
      421 TGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGTCACCGTGCCCGGCCAATATT 480

      977 ACTCCTTTAGTCTCTGGAAAGACTGCGGCT 1006
              |||  | | | | |    || || |
      481 TTGGGTTTTTTTTTTAGTAGAGATGGGGTT 510


LT 14
:743/c
:       CX871743                720 bp    mRNA    linear   EST 03-FEB-2005
NITION  HESC4_60_G05.g1_A037 NIH_MGC_262 Homo sapiens cDNA clone
        IMAGE:7486284 5', mRNA sequence.

ACCESSION   CX871743
VERSION     CX871743.1  GI:58554917
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 720)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: BresaGen, Inc.
             cDNA Library Preparation: Express Genomics, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Laboratory for Genomics and Bioinformatics,
            University of Georgia
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM15813  row: n  column: 10
            Seq primer: JENREV (CAGGAAACAGCTATGACC)
            High quality sequence stop: 720.
FEATURES             Location/Qualifiers
     source          1. .720
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:7486284"
                     /sex="male"
                     /tissue_type="embryonic stem"
                     /cell_type="human embryonic stem cells"
                     /cell_line="BG01"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="NIH_MGC_262"
                     /note="Vector: pExpress-1; Site_1: NotI; Site_2: EcoRV;
                     RNA obtained from human embryonic stem cells Isolated from
                     the inner cell mass of blastocyst stage embryos and
                     differentiated to an early neural progenitor cell type.
                     Cell line id and NIH Registry designation is BG01.
                     Positive for Nestin and Musashi expression. Passage number
                     18. cDNA primed using oligo-dT primer:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3' and cloned into
                     the EcoRV/NotI sites of pExpress-1. This primary library
                     is non-normalized (normalized primary library is
                     NIH_MGC_259). It was constructed by Express Genomics
                     (Frederick, MD). Sequence ends have been trimmed to
                     exclude vector and regions below Phred quality 16. Note:
                     this is a Mammalian Gene Collection library."
ORIGIN

  Query Match            16.0%;  Score 298.6;  DB 9;  Length 720;
  Best Local Similarity  79.1%;  Pred. No. 4.1e-37;
  Matches  406;  Conservative    0;  Mismatches   99;  Indels    8;  Gaps    4;

Qy        484 AGATTGAGCTGCTGCTGCTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTG 543
              || | |  || | | |   ||| ||| ||| |||||  ||||||||||||||||||||
Db        634 AGCTGGGACTACGGTTTGTTTGTTTGTTTGTTTTTTGTTTTTTTTTGAGACAGAGTCT-- 577

Qy        544 CTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCC 603
              | ||||| |||||||||||||||| ||||  ||||||| || ||||||||||||||||||
Db        576 CACTGTCACCCAGGCTGGAGTGCTGTGGTGTGATCTCGGCTCACTGCAACCTCTGCCTCC 517

Qy        604 CGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCC 663
              ||| |||||| ||||||||||||||||||||  ||||||||||||||||||||| ||||
Db        516 CGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGTGCCCGCC 457

Qy        664 ACCACACCCAGC---TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720
```

```
          ||||  ||  ||   ||||||| |||| |||||||||||| || ||||| ||||||||||
      456 ACCATGCCTGGCTAATTTTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGG 397

      721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780
          ||| ||||||   | |||  | |||||||  | |||| ||||||  ||||||  || ||||||
      396 AGTGCAGTGGTGCGGTCTTGGCTCACTGCAAGCTCCGCCTCCCAGGTTCACGCCATTCTC 337

      781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA-TT 839
          |||||||||  |||||||||||||| ||| |||||   |  | || |||||  ||||||| ||
      336 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCAGCTAATTT 277

      840 TTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGG 899
          |||||||||||||||||||| | |||  |  |||| |||| | | |||| || | ||
      276 TTTGTATTTTTAGTAGAGATGGGGTTTCACTGTGTTAGCCAAGATGGTCTCGATCTCCTC 217

      900 ACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCAT 959
          |||||  ||||||| |  |||| ||||||| |||||||||||||| |||||||||| |||
      216 ACCTC--GTGATCCACCCGCCTCGGCCTCCTAAAGTGCTGGGATTACAGGCGTGAGTCAT 159

      960 CGCGCCAGGCCTGAGCTACTCCTTTAGTCTCTG 992
          |||  | |||        | | | | |||| ||
      158 CGCATCCGGCGACCTTCTCCCTTCTTGTCTATG 126


LT 15
5683/c
5       BF876683                537 bp    mRNA    linear   EST 17-JAN-2001
DITION  QV0-ET0149-131100-496-b06 ET0149 Homo sapiens cDNA, mRNA sequence.
SSION   BF876683
CN      BF876683.1  GI:12266722
ORDS    EST.
CE      Homo sapiens (human)
GANISM  Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
        Hominidae; Homo.
ERENCE  1  (bases 1 to 537)
THORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
        Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
        Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
        Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
        O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
        Simpson,A.J.
TLE     Shotgun sequencing of the human transcriptome with ORF expressed
        sequence tags
URNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
UBMED   10737800
ENT     Contact: Simpson A.J.G.
        Laboratory of Cancer Genetics
        Ludwig Institute for Cancer Research
        Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
        Brazil
        Tel: +55-11-2704922
        Fax: +55-11-2707001
        Email: asimpson@ludwig.org.br
        This sequence was derived from the FAPESP/LICR Human Cancer Genome
        Project. This entry can be seen in the following URL
        (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV0&t2=QV0-ET0149-
        131100-496-b06&t3=2000-11-13&t4=1)
        Seq primer: puc 18 forward
        High quality sequence start: 13
        High quality sequence stop: 534.
URES            Location/Qualifiers
 source         1. .537
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /dev_stage="Adult"
                /clone_lib="ET0149"
                /note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
                Site_2: SmaI; A mini-library was made by cloning products
                derived from  ORESTES PCR (U.S. Letters Patent application
                No. 196,716 -  Ludwig Institute for Cancer Research)
```

```
                profiles into  the pUC 18 vector. Reverse transcription of
                tissue mRNA and cDNA amplification were performed under
                low stringency conditions."
ORIGIN

  Query Match             15.9%;  Score 297.2;  DB 2;  Length 537;
  Best Local Similarity   78.4%;  Pred. No. 7.5e-37;
  Matches  373;  Conservative    0;  Mismatches   93;  Indels   10;  Gaps    1;

Qy        519 TCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATC 578
              | |||||||||||||||||| |  |||||| ||||| || ||||||||||||  |||||
Db        521 TTTTTTTTTTGAGACAGAGTTTCACTCTGTTGCCCAAGCCGGAGTGCAGTGGTGTGATC 462

Qy        579 TCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAA 638
              | || |||||||||||| ||||||| ||||| ||| ||||| |  |||||||||||  |
Db        461 GCAGCTCACTGCAACCTCCGCCTCCCAGATTCCAGCAATTCTTCCACCTCAGCCTCCTGA 402

Qy        639 GTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT----------TTTTTTATTTTG 688
              |||||| |||||||||  |  | ||||||  |||||||           ||| ||  |||
Db        401 GTAGCTAGGATTACAGACGTGCACCACCATGCCCAGCTAATTTATTTATTTATTTATTTT 342

Qy        689 GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTG 748
              ||||| |||| |  | |||||||||||||||||| |||||||   |||| ||| ||||
Db        341 GAGACGGAGTTTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCACTGTCTCGGTTTACTG 282

Qy        749 CGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGAT 808
              | |||||  |||||| |||||||||||||||||||| | ||||  ||||||| |||
Db        281 CAACCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCGGCCTCCCAAGTAGCTGGGAC 222

Qy        809 TACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTG 868
              |||||  |||  ||  ||||  |||||||||||||||||||||||||||   | |||||
Db        221 TACAGGCGTGTGCCAACACGCCCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTG 162

Qy        869 CCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTC 928
              |||||||| | | |||  |||  ||| || |||||||||||||| |  |||||||| ||
Db        161 CCATGTTGCCTAGGCTGTTCTCAAACTCCTGACCTCAGGTGATCCACCTGCCTTGGCTTC 102

Qy        929 CCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCCTTT 984
              ||||||||||||||| ||||| |||||||  || || ||| ||   |  |  |||
Db        101 CCAAAGTGCTGGGATTACAGGCATGAGCCACTGCACCTGGCTTGTTTTCTTTTTTT 46


Search completed: June 22, 2006, 01:59:31
Job time : 9233 secs
```

GenCore version 5.1.9

ucleic - nucleic search, using sw model

on: June 21, 2006, 21:46:30 ; Search time 378 Seconds
(without alignments)
9256.544 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ence: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ing table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ched: 1403666 seqs, 935554401 residues

l number of hits satisfying chosen parameters: 2807332

num DB seq length: 0
num DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : Issued_Patents_NA:*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

t o.	Score	% Query Match	Length	DB	ID	Description
1	329.8	17.6	31623	3	US-09-949-016-15945	Sequence 15945, A
2	324.8	17.4	15116	3	US-09-949-016-12354	Sequence 12354, A
3	324.8	17.4	15117	3	US-09-949-016-16260	Sequence 16260, A
4	323.6	17.3	45819	3	US-09-949-002-825	Sequence 825, App
5	323.6	17.3	162450	3	US-09-345-882-1	Sequence 1, Appli
6	323.6	17.3	162450	3	US-10-071-179-1	Sequence 1, Appli
7	323.2	17.3	601	3	US-09-949-002-9649	Sequence 9649, Ap
8	321.8	17.2	50368	3	US-09-949-016-13256	Sequence 13256, A
9	320.6	17.1	129554	3	US-09-949-002-765	Sequence 765, App
10	319.8	17.1	13426	3	US-09-949-016-13683	Sequence 13683, A
11	316	16.9	83851	3	US-09-949-016-13847	Sequence 13847, A
12	315.8	16.9	7481	3	US-09-949-016-14492	Sequence 14492, A
13	315.8	16.9	21308	3	US-09-949-002-584	Sequence 584, App
14	315.8	16.9	21407	3	US-09-949-002-784	Sequence 784, App
15	312.8	16.7	40742	3	US-09-949-016-11751	Sequence 11751, A
16	312.8	16.7	40747	3	US-09-949-016-13097	Sequence 13097, A
17	312.8	16.7	90472	3	US-09-949-016-14038	Sequence 14038, A
18	312.8	16.7	99014	4	US-09-880-107-3428	Sequence 3428, Ap
19	311.8	16.7	102053	3	US-09-949-016-13025	Sequence 13025, A
20	308.6	16.5	81001	3	US-09-750-580-1	Sequence 1, Appli
21	307.6	16.4	21372	3	US-09-949-016-16941	Sequence 16941, A
22	307.4	16.4	9822	3	US-09-949-016-15697	Sequence 15697, A
23	307	16.4	6569	3	US-09-949-016-17534	Sequence 17534, A
24	306.8	16.4	45427	3	US-09-949-016-16243	Sequence 16243, A
c 25	306.6	16.4	71574	3	US-09-949-016-15580	Sequence 15580, A
26	305.4	16.3	112112	3	US-09-949-016-15639	Sequence 15639, A
c 27	304.8	16.3	601	3	US-09-949-016-63238	Sequence 63238, A
28	304.8	16.3	601	3	US-09-949-016-174955	Sequence 174955,
c 29	304.8	16.3	43577	3	US-09-949-016-16694	Sequence 16694, A
30	304.6	16.3	28806	3	US-09-949-016-13217	Sequence 13217, A
31	303.6	16.2	13677	3	US-09-949-016-16817	Sequence 16817, A
32	303.6	16.2	80355	3	US-09-949-016-12735	Sequence 12735, A
33	303.6	16.2	80357	3	US-09-949-016-13572	Sequence 13572, A
34	303.6	16.2	134890	3	US-09-949-016-15602	Sequence 15602, A
c 35	303.4	16.2	47030	3	US-09-949-016-13037	Sequence 13037, A
c 36	303.4	16.2	47030	3	US-09-949-016-13038	Sequence 13038, A
c 37	303.4	16.2	47030	3	US-09-949-016-15039	Sequence 15039, A
c 38	303.4	16.2	47030	3	US-09-949-016-15040	Sequence 15040, A
39	303.2	16.2	16572	3	US-09-949-016-12409	Sequence 12409, A
40	303.2	16.2	16578	3	US-09-949-016-14680	Sequence 14680, A
41	303.2	16.2	44342	3	US-09-949-016-12661	Sequence 12661, A
c 42	302.8	16.2	33152	3	US-09-949-016-16719	Sequence 16719, A
43	302.2	16.2	105413	3	US-10-427-923-3	Sequence 3, Appli
44	302.2	16.2	112219	3	US-09-949-016-12453	Sequence 12453, A
45	302.2	16.2	112222	3	US-09-949-016-14324	Sequence 14324, A

ALIGNMENTS

RESULT 1
US-09-949-016-15945
; Sequence 15945, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15945
; LENGTH: 31623
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15945

Query Match 17.6%; Score 329.8; DB 3; Length 31623;
Best Local Similarity 81.0%; Pred. No. 9.2e-81;
Matches 396; Conservative 0; Mismatches 92; Indels 1; Gaps 1;

```
Qy        498 CTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGG 557
              ||| |||||||||||| ||||| ||||||| ||||||| |||| ||||||||||||||
Db       9596 CTGTTTTTTTTTTTTTTTTTTTTTTTTTTAAGAGACAGGGTCTCACTCTGTCGCCCAGG 9655

Qy        558 CTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGAT 617
              |||||||| |||| ||||||| || ||||| ||||| ||||||| || |||||| |||
Db       9656 CTGGAGTGTGGTGGTGTGATCTCAGCTCACTGGAACCTTTGCCTCCTGGGTTCAAGGGAT 9715

Qy        618 TCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTT 677
              ||||||||||||||||    ||| ||||||||||||||  |  ||||||||||||||||
Db       9716 TCTCCTGCCTCAGCCTCATGAGTGGCTGGGATTACAGGCACGAGCCACCACACCCAGCTA 9775

Qy        678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
              |||| |||| |||| |||||| || ||||||||||||||||||| | ||||||  |||
Db       9776 ATTTTTTTTTTGAGATGGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCGGTGGCACCATC 9835

Qy        738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
```

```
      |||| ||||||| |||||  |||||| ||||||||| |||||||||||||||| |||| ||
 9836 TCAGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGA 9895

  798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
      ||||| ||| |||||  ||| | | ||| ||   |||||||||| ||| ||||||||||||
 9896 ATAGCTGGGACTACAGGTGTGTATCACCATGTCTGGCTAATTTTAGTA-TTTTAGTAGAG 9954

  858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
      || | |||  |||||||||||| ||| |||| |||| | |||||||| ||||||| |
 9955 CTGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTTCTGACCTCAAGTGATCCACCC 10014

  918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTA 977
      |||||||||||||||||| ||||||| |||||||||||| | |    |||||    |
10015 ACCTTGGCCTCCCAAAGTGGTGGGATTACAGGCGTGAGCCACCTCAGTTGGCCTCCATTT 10074

  978 CTCCTTTAG 986
      ||  || ||
10075 CTTATTGAG 10083
```

```
;T 2
)-949-016-12354
]uence 12354, Application US/09949016
:ent No. 6812339
:ERAL INFORMATION:
'PLICANT: VENTER, J. Craig et al.
:TLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
:TLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
:LE REFERENCE: CL001307
]RRENT APPLICATION NUMBER: US/09/949,016
]RRENT FILING DATE:  2000-04-14
'IOR APPLICATION NUMBER: 60/241,755
'IOR FILING DATE: 2000-10-20
'IOR APPLICATION NUMBER: 60/237,768
'IOR FILING DATE: 2000-10-03
'IOR APPLICATION NUMBER: 60/231,498
'IOR FILING DATE: 2000-09-08
'IMBER OF SEQ ID NOS: 207012
)FTWARE: FastSEQ for Windows Version 4.0
'Q ID NO 12354
.ENGTH: 15116
'YPE: DNA
)RGANISM: Human
)-949-016-12354

:ry Match            17.4%;  Score 324.8;  DB 3;  Length 15116;
;t Local Similarity  77.5%;  Pred. No. 1.5e-79;
:ches  409;  Conservative    0;  Mismatches  112;  Indels    7;  Gaps    1;

  502 TTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
      |||||||||||| ||| |  | |||||||||||||||||  |||| || |||||| |||
 5037 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 5096

  562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
      | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
 5097 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 5156

  622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
      ||||||||||||||  |||||||||| ||||||||| | |||| || ||| ||
 5157 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 5216

  676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
       || |||| |||| ||||  ||| ||  | ||||  |||||||||||| || ||||||||
 5217 TTTCTTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 5276

  735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
      ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||| ||||
 5277 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 5336

  795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
      |||||||| |||||||||  |  | || |||||| | |||||||||| ||||||||||
 5337 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 5396
```

```
Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||
Db       5397 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 5456

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
              |  | |||||| ||| |||||||||||||| |||| |||||||| ||||||  |||||
Db       5457 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 5516

Qy        975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
              | ||   ||  | | |    |||| |   |     |  | | | |||
Db       5517 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 5564


RESULT 3
US-09-949-016-16260
; Sequence 16260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16260
;   LENGTH: 15117
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16260

  Query Match            17.4%;  Score 324.8;  DB 3;  Length 15117;
  Best Local Similarity  77.5%;  Pred. No. 1.5e-79;
  Matches  409;  Conservative    0;  Mismatches  112;  Indels    7;  Gaps    1;

Qy        502 TTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
              |||||||||||| ||| |  | |||||||||||||||||  |||| || |||||| |||
Db       5037 TTTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 5096

Qy        562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
              | ||||||||  ||||||  || ||| |||||||||||||||| | |||||| |||||||
Db       5097 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 5156

Qy        622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
              ||||||||||||||  |||||||||| ||||||||| | |||| || ||| ||
Db       5157 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 5216

Qy        676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
               || |||| |||| ||||  ||| ||  | ||||  |||||||||||| || ||||||||
Db       5217 TTTCTTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 5276

Qy        735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
              ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||| ||||
Db       5277 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 5336

Qy        795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
              |||||||| |||||||||  |  | || |||||| | |||||||||| ||||||||||
Db       5337 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 5396

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||
Db       5397 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 5456

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
              |  | |||||| ||| |||||||||||||| |||| |||||||| ||||||  |||||
Db       5457 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 5516
```

```
      975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
          | ||   ||  | | |    |||| |   |     | | | | |||
     5517 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 5564


LT 4
9-949-002-825/c
quence 825, Application US/09949002
tent No. 6900016
NERAL INFORMATION:
PPLICANT: VENTER, J. Craig et al.
ITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
ITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
ITLE OF INVENTION:  AND USES THEREOF
ILE REFERENCE: CL000790
URRENT APPLICATION NUMBER: US/09/949,002
URRENT FILING DATE:  2000-01-28
RIOR APPLICATION NUMBER: 60/231,401
RIOR FILING DATE: 2000-09-08
UMBER OF SEQ ID NOS: 10823
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 825
ENGTH: 45819
YPE: DNA
RGANISM: Human
9-949-002-825

ery Match             17.3%;  Score 323.6;  DB 3;  Length 45819;
t Local Similarity   82.9%;  Pred. No. 5.9e-79;
tches  383;  Conservative    0;  Mismatches   74;  Indels    5;  Gaps    1;

      517 TTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGA 576
          ||| |||||||||||||  ||||||  |||||||||| ||| |||||||||||||    |
    23998 TTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCGCCTAGGGTGGAGTGCAGTGGCGCAA 23939

      577 TCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCC 636
          |||| || |||||||||||||| |||||||||  ||||||  |||||||||||||||||||||
    23938 TCTCAGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCC 23879

      637 AAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTTTATTTTGGAG 691
          |||||||||||| |||| |  || |||||  | ||||||     ||||||| |||| |||
    23878 AAGTAGCTGGGACTACAAGCACATGCCACTGCGCCCAGCTAATTTTTTTTTTTTTTTGAG 23819

      692 ACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGA 751
          |  ||||||  | |||||  || ||| ||.||| |||||||   ||||||| |||||||  |
    23818 ATGGAGTCTCACTCTGTCGCCTAGGGTGGAGTGCAGTGGCGCAATCTCAGCTCACTGCAA 23759

      752 CCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTAC 811
          ||||| ||||||| |||||||  |||||||||||||| |||||||||||||| |||||||
    23758 CCTCCGCCTCCCGAGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTAC 23699

      812 AGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCA 871
          ||  |  |||| |||||  |||||||||||||||||||||||||||||  |    || ||||
    23698 AGGTGCCCACCACCACGCCCAGCTAATTTTTGTATTTTTAGTAGAGACGGTGCTTCGCCA 23639

      872 TGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCA 931
          ||||||||| ||| |||||| ||  | ||||||||||||||| |  | || |||||||||
    23638 TGTTGGCCAGGCTGGTCTGGCACTACTGACCTCAGGTGATCCACCCGTCTCGGCCTCCCA 23579

      932 AAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGA 973
          |||||||||||||| ||||| ||||||| |||||| |||| ||
    23578 AAGTGCTGGGATTACAGGCATGAGCCAGCGCGCCCGGCCAGA 23537


LT 5
9-345-882-1/c
quence 1, Application US/09345882
tent No. 6399373
NERAL INFORMATION:
PPLICANT: Bougueleret, Lydie
ITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
ITLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
;  FILE REFERENCE: GENSET.031A
;  CURRENT APPLICATION NUMBER: US/09/345,882
;  CURRENT FILING DATE:  1999-06-30
;  PRIOR APPLICATION NUMBER: US 60/091,315
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: US 60/111,909
;  PRIOR FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 140
;  SOFTWARE: Patent.pm
; SEQ ID NO 1
;   LENGTH: 162450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72794
;   OTHER INFORMATION: 5-124-273  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 88073
;   OTHER INFORMATION: 5-127-261  :  polymorphic  base  A  or  C
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 90842
;   OTHER INFORMATION: 99-1437-325  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 93714
;   OTHER INFORMATION: 5-128-60  :  polymorphic  base  deletion  of  GT
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97122
;   OTHER INFORMATION: 99-1442-224  :  polymorphic  base  G  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97152
;   OTHER INFORMATION: 5-129-144  :  polymorphic  base  deletion  of  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99098
;   OTHER INFORMATION: 5-130-257  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99117
;   OTHER INFORMATION: 5-130-276  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103806
;   OTHER INFORMATION: 5-131-395  :  polymorphic  base  A  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106940
;   OTHER INFORMATION: 5-133-375  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108106
;   OTHER INFORMATION: 5-135-155  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108149
;   OTHER INFORMATION: 5-135-198  :  polymorphic  base  insertion  of  GTTT
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108308
;   OTHER INFORMATION: 5-135-357  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108471
;   OTHER INFORMATION: 5-136-174  :  polymorphic  base  C  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 134134
;   OTHER INFORMATION: 5-140-120  :  polymorphic  base  C  or  T
```

```
'EATURE:
'AME/KEY: allele
.OCATION: 134362
)THER INFORMATION: 5-140-348  :  polymorphic  base  insertion  of  A
'EATURE:
'AME/KEY: allele
.OCATION: 134374
)THER INFORMATION: 5-140-361  :  polymorphic  base  insertion  of  CA
'EATURE:
'AME/KEY: allele
.OCATION: 146328
)THER INFORMATION: 5-143-84  :  polymorphic  base  A  or  G
'EATURE:
'AME/KEY: allele
.OCATION: 146345
)THER INFORMATION: 5-143-101  :  polymorphic  base  A  or  C
'EATURE:
'AME/KEY: allele
.OCATION: 150329
)THER INFORMATION: 5-145-24  :  polymorphic  base  A  or  G
'EATURE:
'AME/KEY: allele
.OCATION: 160031
)THER INFORMATION: 5-148-352  :  polymorphic  base  G  or  T
'EATURE:
'AME/KEY: allele
.OCATION: 72771..72817
)THER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
'EATURE:
'AME/KEY: allele
.OCATION: 72771..72817
)THER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
'EATURE:
'AME/KEY: allele
.OCATION: 88050..88096
)THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
'EATURE:
'AME/KEY: allele
.OCATION: 88050..88096
)THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
'EATURE:
'AME/KEY: allele
.OCATION: 90819..90865
)THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
'EATURE:
'AME/KEY: allele
.OCATION: 90819..90865
)THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
'EATURE:
'AME/KEY: allele
.OCATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
'EATURE:
'AME/KEY: allele
.OCATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
'EATURE:
'AME/KEY: allele
.OCATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
'EATURE:
'AME/KEY: allele
.OCATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
'EATURE:
'AME/KEY: allele
.OCATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
'EATURE:
'AME/KEY: allele
.OCATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
'EATURE:
```

```
;  NAME/KEY: allele
;  LOCATION: 99075..99121
;  OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 99075..99121
;  OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 99094..99140
;  OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 99094..99140
;  OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 103783..103828
;  OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 103783..103828
;  OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 106918..106966
;  OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 106918..106966
;  OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108084..108130
;  OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108084..108130
;  OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108127..108177
;  OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108127..108177
;  OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
;  FEATURE:

  Query Match             17.3%;  Score 323.6;  DB 3;  Length 162450;
  Best Local Similarity   81.2%;  Pred. No. 1.1e-78;
  Matches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              |  ||  ||  |  |   ||  |||||||  ||||   ||  ||||||||||||||||||||||
Db     102008 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
              ||||||||||||||||||||||||||  |  |||||  ||  ||||||||||||  ||  |||  ||
Db     101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

Qy        608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
              |||||||  ||  ||||||||||||||||  ||||||||||||  |||||  |  |||||||
Db     101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

Qy        668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
              |  |||||||  ||||||  ||||  |||  ||||||  ||  ||||||||||||||||||  |
Db     101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

Qy        726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
              ||||||  |||||||  |  |||||||  |  ||||  ||||||||||  |  |  ||||||||||
Db     101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709
```

```
     786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
         |||| |||||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
  101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

     844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
         | |||||||||||||    | |||  ||||||| |||| | | |||| || | || |||||
  101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

     904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
         |||||||| ||  |||||||||||||||||||||||||| |||||||||||| ||||
  101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

     964 CCAGGCCTGAGCTA 977
         ||  |||  | | |
  101528 CCCAGCCACACCCA 101515
```

```
LT 6
.0-071-179-1/c
equence 1, Application US/10071179
tent No. 6908988
ENERAL INFORMATION:
PPLICANT: Bougueleret, Lydie
ITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
ITLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
ILE REFERENCE: GENSET.031A
URRENT APPLICATION NUMBER: US/10/071,179
URRENT FILING DATE:  2002-02-07
RIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/345,882
RIOR FILING DATE: EARLIER FILING DATE: 1999-06-30
RIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/091,315
RIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
RIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/111,909
RIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
UMBER OF SEQ ID NOS: 140
OFTWARE: Patent.pm
Q ID NO 1
ENGTH: 162450
YPE: DNA
RGANISM: Homo sapiens
EATURE:
AME/KEY: allele
OCATION: 72794
OTHER INFORMATION: 5-124-273  :  polymorphic  base  A  or  G
EATURE:
AME/KEY: allele
OCATION: 88073
OTHER INFORMATION: 5-127-261  :  polymorphic  base  A  or  C
EATURE:
AME/KEY: allele
OCATION: 90842
OTHER INFORMATION: 99-1437-325  :  polymorphic  base  A  or  G
EATURE:
AME/KEY: allele
OCATION: 93714
OTHER INFORMATION: 5-128-60  :  polymorphic  base  deletion  of  GT
EATURE:
AME/KEY: allele
OCATION: 97122
OTHER INFORMATION: 99-1442-224  :  polymorphic  base  G  or  T
EATURE:
AME/KEY: allele
OCATION: 97152
OTHER INFORMATION: 5-129-144  :  polymorphic  base  deletion  of  T
EATURE:
AME/KEY: allele
OCATION: 99098
OTHER INFORMATION: 5-130-257  :  polymorphic  base  A  or  G
EATURE:
AME/KEY: allele
OCATION: 99117
OTHER INFORMATION: 5-130-276  :  polymorphic  base  A  or  G
EATURE:
```

```
;  NAME/KEY: allele
;  LOCATION: 103806
;  OTHER INFORMATION: 5-131-395  :  polymorphic  base  A  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 106940
;  OTHER INFORMATION: 5-133-375  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108106
;  OTHER INFORMATION: 5-135-155  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108149
;  OTHER INFORMATION: 5-135-198  :  polymorphic  base  insertion  of  GTTT
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108308
;  OTHER INFORMATION: 5-135-357  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108471
;  OTHER INFORMATION: 5-136-174  :  polymorphic  base  C  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134134
;  OTHER INFORMATION: 5-140-120  :  polymorphic  base  C  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134362
;  OTHER INFORMATION: 5-140-348  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134374
;  OTHER INFORMATION: 5-140-361  :  polymorphic  base  insertion  of  CA
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 146328
;  OTHER INFORMATION: 5-143-84  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 146345
;  OTHER INFORMATION: 5-143-101  :  polymorphic  base  A  or  C
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 150329
;  OTHER INFORMATION: 5-145-24  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 160031
;  OTHER INFORMATION: 5-148-352  :  polymorphic  base  G  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 72771..72817
;  OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 72771..72817
;  OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 88050..88096
;  OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 88050..88096
;  OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 90819..90865
;  OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
;  FEATURE:
;  NAME/KEY: allele
```

```
)CATION: 90819..90865
)THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
'EATURE:
!AME/KEY: allele
)CATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
'EATURE:
!AME/KEY: allele
)CATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
'EATURE:
!AME/KEY: allele
)CATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
'EATURE:
!AME/KEY: allele
)CATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
'EATURE:
!AME/KEY: allele
)CATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
'EATURE:
!AME/KEY: allele
)CATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
'EATURE:
!AME/KEY: allele
)CATION: 99075..99121
)THER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
'EATURE:
!AME/KEY: allele
)CATION: 99075..99121
)THER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
'EATURE:
!AME/KEY: allele
)CATION: 99094..99140
)THER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
'EATURE:
!AME/KEY: allele
)CATION: 99094..99140
)THER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
'EATURE:
!AME/KEY: allele
)CATION: 103783..103828
)THER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
'EATURE:
!AME/KEY: allele
)CATION: 103783..103828
)THER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
'EATURE:
!AME/KEY: allele
)CATION: 106918..106966
)THER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
'EATURE:
!AME/KEY: allele
)CATION: 106918..106966
)THER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
'EATURE:
!AME/KEY: allele
)CATION: 108084..108130
)THER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
'EATURE:
!AME/KEY: allele
)CATION: 108084..108130
'THER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
'EATURE:
!AME/KEY: allele
)CATION: 108127..108177
)THER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
'EATURE:
!AME/KEY: allele
)CATION: 108127..108177
```

```
  Query Match            17.3%;  Score 323.6;  DB 3;  Length 162450;
  Best Local Similarity  81.2%;  Pred. No. 1.1e-78;
  Matches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              | ||  || |  |   || ||||||| ||||   || ||||||||||||||||||||||
Db     102008 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
              |||||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
Db     101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

Qy        608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
              ||||||| || || |||||||||||||| |||||||||||| |||||  |  ||||||
Db     101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

Qy        668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
              | |||||||  ||||||  ||||  |||  ||||||  || ||||||||||||||||||| |
Db     101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

Qy        726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
              ||||||  |||||||  | |||||||  | ||||  ||||||||||| |  |  |||||||||||
Db     101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709

Qy        786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
              ||||  |||||||||||||| |||  |||||  ||  ||||  |||  |  |  ||||||||||  |
Db     101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

Qy        844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
              | ||||||||||||||   | |||  |||||||  ||||  |  |  ||||  ||  |  ||  |||||
Db     101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

Qy        904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
              |||||||||  ||  ||||||||||||||||||||||||||||||||  ||||||||||||  ||||
Db     101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

Qy        964 CCAGGCCTGAGCTA 977
              ||  |||  |  | |
Db     101528 CCCAGCCACACCCA 101515


RESULT 7
US-09-949-002-9649/c
; Sequence 9649, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9649
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-9649

  Query Match            17.3%;  Score 323.2;  DB 3;  Length 601;
  Best Local Similarity  82.7%;  Pred. No. 8.2e-80;
  Matches  382;  Conservative    1;  Mismatches   74;  Indels    5;  Gaps    1;

Qy        517 TTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGA 576
              |||  |||||||||||||  ||||||  ||||||||||  ||| ||||||||||||||   |
Db        511 TTTTTTTTTTTTTGAGATGGAGTCTCACTCTGTCGCCTAGGGTGGAGTGCAGTGGCGCAA 452

Qy        577 TCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCC 636
```

```
     |||| || |||||||||||| ||||||||| |||||| |||||||||||||||||||||
  451 TCTCAGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGAGATTCTCCTGCCTCAGCCTCCC 392

  637 AAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTTTATTTTGGAG 691
      |||||||||||| |||| |  || |||||  | ||||||      |||||| |||| |||
  391 AAGTAGCTGGGACTACAAGCACATGCCACTGCGCCCAGCTAATTTTTTTTTTTTTTTGAG 332

  692 ACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGA 751
      |  ||||||  | ||||| || ||| ||||: |||||||   ||||||| ||||||| |
  331 ATGGAGTCTCACTCTGTCGCCTAGGGTGGARTGCAGTGGCGCAATCTCAGCTCACTGCAA 272

  752 CCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTAC 811
      ||||| ||||||| ||||||| |||||||||||||| |||||||||||| |||||||
  271 CCTCCGCCTCCCGAGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTAC 212

  812 AGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCA 871
      ||  |  |||| |||||  |||||||||||||||||||||||||||| |    || ||||
  211 AGGTGCCCACCACCACGCCCAGCTAATTTTTGTATTTTTAGTAGAGACGGTGCTTCGCCA 152

  872 TGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCA 931
      |||||||| ||| |||||| ||  | ||||||||||||||| |  | || |||||||||
  151 TGTTGGCCAGGCTGGTCTGGCACTACTGACCTCAGGTGATCCACCCGTCTCGGCCTCCCA 92

  932 AAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGA 973
      |||||||||||| ||||| ||||||| |||||| |||| ||
   91 AAGTGCTGGGATTACAGGCATGAGCCAGCGCGCCCGGCCAGA 50
```

LT 8
9-949-016-13256
quence 13256, Application US/09949016
ent No. 6812339
NERAL INFORMATION:
PPLICANT: VENTER, J. Craig et al.
ITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
ITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
ILE REFERENCE: CL001307
URRENT APPLICATION NUMBER: US/09/949,016
URRENT FILING DATE: 2000-04-14
RIOR APPLICATION NUMBER: 60/241,755
RIOR FILING DATE: 2000-10-20
RIOR APPLICATION NUMBER: 60/237,768
RIOR FILING DATE: 2000-10-03
RIOR APPLICATION NUMBER: 60/231,498
RIOR FILING DATE: 2000-09-08
UMBER OF SEQ ID NOS: 207012
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 13256
LENGTH: 50368
TYPE: DNA
ORGANISM: Human
9-949-016-13256

ery Match 17.2%; Score 321.8; DB 3; Length 50368;
t Local Similarity 81.8%; Pred. No. 2e-78;
tches 387; Conservative 0; Mismatches 77; Indels 9; Gaps 1;

```
      507 TTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGC 566
          ||  ||| ||||| || |||||||||||||| |||||||| ||| ||||||||||||||||
    48915 TTGATTTTTTTTTTTTCTTTTTTGAGACAGGGTCTTGCTTTGTTGCCCAGGCTGGAGTGC 48974

      567 AGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCC 626
          ||||||| |||||| || |||||||||||||| |||| | |||||| ||||||||||||
    48975 AGTGGCACGATCTCGGCTCACTGCAACCTCTGCATCCCAGGTTCAAGTGATTCTCCTGCC 49034

      627 TCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---------TT 677
          | |||||||  |||||||||||||||||||      | ||  |||||  ||         ||
    49035 TTAGCCTCCTGAGTAGCTGGGATTACAGGCATGTGACAAGACACCTGGCTAATTTTTTTT 49094

      678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
          ||||| |||| ||||||||||||    |||||||||||||||| ||| || ||||||||||
    49095 TTTTTTTTTTTGAGACAGAGTCTCCTTCTGTCACCCAGGCTGAAGTGCAATGGCATGATC 49154
```

```
Qy        738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
              || |||||||||  |||||||||||||||  ||||  |||||||||||||||||||| ||||  |
Db      49155 TCGGTTCACTGCAACCTCCACCTCCCAGGTTTGAGCAATTCTCCTGCCTCAGCCTCCCTA 49214

Qy        798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
              ||||||  |||||||||     ||||| ||| ||   ||||||||||||||||||||||||||
Db      49215 GTAGCTGGGATTACAGGTACGCACCACCATGTCTGGCTAATTTTTGTATTTTTAGTAGAG 49274

Qy        858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
              ||| | ||| |||||||||||||| ||| ||||  ||| || ||| |||||||||| |
Db      49275 ATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGACTTCAGGTGATCCACCT 49334

Qy        918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
              | || |||||| |||||| |||||||| ||||| |||||||  ||||| ||||
Db      49335 GTCTCGGCCTCTCAAAGTTCTGGGATTACAGGCATGAGCCACTGCGCCTGGCC 49387
```

RESULT 9
US-09-949-002-765
; Sequence 765, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 129554
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(129554)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-765

Query Match 17.1%; Score 320.6; DB 3; Length 129554;
Best Local Similarity 81.4%; Pred. No. 6.9e-78;
Matches 385; Conservative 0; Mismatches 84; Indels 4; Gaps 1;

```
Qy        515 TTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCAT 574
              ||||| |||||||||  || | || ||  |||||| ||||||||||||||||| |||
Db      35451 TTTTTTTTTTTTTTTTGAGATAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGTGC 35510

Qy        575 GATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTC 634
              ||||| || |||||||||||||||||||| || |||||||||||||||||||||| | ||
Db      35511 GATCTTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCACCTTC 35570

Qy        635 CCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----TTTTTTTATTTTGGA 690
              || ||| | |||||||||||||||| | ||||||||||| ||    ||||||| || | ||
Db      35571 CCGAGTTGTTGGGATTACAGGTGCCCACCACCACACCCGGCTGATTTTTTTTTTTCTTGA 35630

Qy        691 GACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCG 750
              ||| ||||||||| ||||| ||||||||| ||| | |||||  |||||| | | |||||
Db      35631 GACGGAGTCTTGCTCTGTCGCCCAGGCTGAAGTGCGGTGGCGCGATCTCTGCTTACTGCA 35690

Qy        751 ACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTA 810
              | |||  |||||| |||||||||  |||||||||||||||  |||||||||||||| ||||||
Db      35691 AACTCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTA 35750

Qy        811 CAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCC 870
              |||  |||  || ||| |  ||||||||||||||||||||||||| ||||| | |  ||||||
Db      35751 CAGGTGTGTGCCACCATGCCCAGCTAATTTTTGTATTTTTAGGAGAGACGGGGCTTTGCC 35810

Qy        871 ATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCC 930
              ||||||| || ||| |||| |||| || ||||||||||||||| ||  |||| ||||||||
```

```
    35811 ATGTTGGTCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCTGCCTGCCTCGGCCTCCC 35870

      931 AAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCCTT 983
          |||||||||||||| |||| |||||||| |  ||| |||||  |  | |  ||
    35871 AAAGTGCTGGGATTTCAGGTGTGAGCCACCATGCCCGGCCTTTGGAAATTTTT 35923


.T 10
)-949-016-13683
;uence 13683, Application US/09949016
ent No. 6812339
:ERAL INFORMATION:
:PLICANT: VENTER, J. Craig et al.
:TLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
:TLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
:LE REFERENCE: CL001307
)RRENT APPLICATION NUMBER: US/09/949,016
)RRENT FILING DATE:  2000-04-14
:IOR APPLICATION NUMBER: 60/241,755
:IOR FILING DATE: 2000-10-20
:IOR APPLICATION NUMBER: 60/237,768
:IOR FILING DATE: 2000-10-03
:IOR APPLICATION NUMBER: 60/231,498
:IOR FILING DATE: 2000-09-08
)MBER OF SEQ ID NOS: 207012
)FTWARE: FastSEQ for Windows Version 4.0
) ID NO 13683
.ENGTH: 13426
:YPE: DNA
)RGANISM: Human
:EATURE:
)AME/KEY: misc_feature
)CATION: (1)...(13426)
:THER INFORMATION: n = A,T,C or G
)-949-016-13683

:ry Match            17.1%;  Score 319.8;  DB 3;  Length 13426;
:t Local Similarity   82.9%;  Pred. No. 3.6e-78;
:ches  389;  Conservative    0;  Mismatches   77;  Indels    3;  Gaps    2;

      505 TTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT-GTCGCCCAGGCTGGAG 563
          |||||||||| ||||| |||||||||| |||||||| | ||||| || ||||||||||||
     6850 TTTTTTTTTTTTTTTTTTTTTTTTGGTACAGAGTTTCGCTCTGGTTGCCCAGGCTGGAG 6909

      564 TGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCT 623
          |||| ||||| ||||| || |||||||| |||||||||||||||| ||||||||||||||||||
     6910 TGCAATGGCACGATCTTGGCTCACTGCAGCCTCTGCCTCCCGGGTTCAAGCGATTCTCCT 6969

      624 GCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT--TTTTT 681
          |||||||||||  || ||||||||||||||     | ||||||  ||| |||  |||||
     6970 GCCTCAGCCTCCGGAGCAGCTGGGATTACAGACATGCACCACCATGCCCGGCTAATTTTT 7029

      682 TATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAG 741
          | |||| ||||  ||||||  |  |||||||||||||||||| ||||||||  ||||| |
     7030 TTTTTTCGAGATGGAGTCTCACTGTGTCACCCAGGCTGGAGTGCAGTGGCACAATCTCGG 7089

      742 TTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAG 801
          |||||||| |||||  |||||||||||||||||| |||||||||||||||| ||||||||||
     7090 CTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAG 7149

      802 CTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGC 861
          || ||| |||||  |    || ||||   |||||||||||||| |||||||||||||| |
     7150 CTGGGACTACAGGTGCCTGCCACCACACCCAGCTAATTTTTGTACTTTTAGTAGAGACGG 7209

      862 GCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCT 921
          | |||| |||||||||||| ||| |||| |||| | |||||||||||||||||  |   |||
     7210 GGTTTTACCATGTTGGCCAGGCTGGTCTTGAACTTCTGACCTCAGGTGATCCACCCACCT 7269

      922 TGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          || ||||||||||||||||||| |||||||||||||| || ||| ||||
     7270 CGGTCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCC 7318
```

```
RESULT 11
US-09-949-016-13847/c
; Sequence 13847, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13847
;   LENGTH: 83851
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-13847

  Query Match            16.9%;  Score 316;  DB 3;  Length 83851;
  Best Local Similarity   73.6%;  Pred. No. 1.1e-76;
  Matches  443;  Conservative    0;  Mismatches  155;  Indels    4;  Gaps    3;

Qy        371 ATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTTGGGTCTCAAT 430
              || | | ||||  ||  | |      ||      | | | ||||     || |  |   |
Db      45594 ATACGAAGAGGGTGAGGAGACACACAGGGCCAGCTAGATATCTTTCAGATGAGATTTTTT 45535

Qy        431 CAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTGCAGAGATTGA 490
              |    ||| |  |||  || ||  |||| | ||  |    || |    |  || |
Db      45534 AAAAATTTCAAGCAGAATTGGTAAGCTGCAGGGTCAAGGCTTCAAATGGACTCAGGGTTT 45475

Qy        491 GCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTC 550
              |||   ||  ||||| |||||| |||||    | |||||||||  ||||||  ||||||
Db      45474 GCT--CTCTTTTTTTTCTTTTTTTTTTTTTCCTGTTTTTGAGATGGAGTCTCACTCTGTT 45417

Qy        551 GCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTC 610
              |||||| ||||||||||||||||   ||||| || ||||||| |||||||||| || |||
Db      45416 GCCCAGACTGGAGTGCAGTGGCGCAATCTCAGCTCACTGCAGCCTCTGCCTCCTGGGTTC 45357

Qy        611 AAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACAC 670
              |||||||||||||||    | ||||||||||||||||||||||||||    | |||||  |
Db      45356 CAGCGATTCTCCTGCCCTGGTCTCCCAAGTAGCTGGGATTACAGGCATGCACCACCGTGC 45297

Qy        671 CCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGG 730
              |  |||  |||| |||| || || |||||||||  |||| ||||||||  |||| ||||||
Db      45296 CTGGCTAATTTT-TTTTTGAAACCGAGTCTTGCTTTGTCGCCCAGGCTAGAGTGCAGTGG 45238

Qy        731 CATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGT 790
              || |||||||| |||||||  ||||| || |||  |||||||||  |||||||||  ||||||
Db      45237 CAGGATCTCAGCTCACTGCAACCTCTACTTCCTGGGTTCAAGTTATTCTCCTGTCTCAGT 45178

Qy        791 CTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTT 850
              |||  |||||||| |||||||| || ||||||  |||||  ||||||||||||||||||
Db      45177 CTCTGGAGTAGCTGGGATTACA-AATTGCACCACCACGCCCAGCTAATTTTTGTATTTTT 45119

Qy        851 AGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGA 910
              ||||  || || | |||  |||||||||||| ||| ||||   ||| || ||||||||||||
Db      45118 AGTAAAGCTGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTGA 45059

Qy        911 TCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
              || ||  |||| |||||||||||| |||| |||| ||||| |||||||  ||||| ||||
Db      45058 TCTGCCAGCCTCGGCCTCCCAAAGCGCTGAGATTACAGGCATGAGCCACTGCGCCTGGCC 44999

Qy        971 TG 972
              ||
Db      44998 TG 44997
```

```
LT 12
9-949-016-14492/c
quence 14492, Application US/09949016
tent No. 6812339
NERAL INFORMATION:
PPLICANT: VENTER, J. Craig et al.
TLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
LE REFERENCE: CL001307
URRENT APPLICATION NUMBER: US/09/949,016
URRENT FILING DATE:  2000-04-14
RIOR APPLICATION NUMBER: 60/241,755
RIOR FILING DATE: 2000-10-20
RIOR APPLICATION NUMBER: 60/237,768
RIOR FILING DATE: 2000-10-03
RIOR APPLICATION NUMBER: 60/231,498
RIOR FILING DATE: 2000-09-08
UMBER OF SEQ ID NOS: 207012
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 14492
ENGTH: 7481
YPE: DNA
RGANISM: Human
9-949-016-14492

ery Match             16.9%;  Score 315.8;  DB 3;  Length 7481;
at Local Similarity   76.5%;  Pred. No. 3.4e-77;
tches  400;  Conservative    0;  Mismatches  122;  Indels    1;  Gaps    1;

      455 TCTGCTTGCTCCTGAGGCCAGAGCTGCAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTC 514
          || || |||  | |   ||  |  ||| | ||||       ||   |          |
     7437 TCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCACGACCGGCC 7378

      515 TTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCAT 574
          |||||||||| |||||||||||||||||  |||||||||||||||||||||||||||||
     7377 TTTTTCTTTTCTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCGC 7318

      575 GATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTC 634
          |||||  ||  |||||||  |||||||||| |||  ||||||  ||||||  ||||||||||||
     7317 AATCTCGGCTCACTGCATCCTCTGCCTCTCGGGTTCAAGTGATTCTTGTGCCTCAGCCTC 7258

      635 CCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT-TTTTTTATTTTGGAGAC 693
          |  ||||||||  ||  ||||  |  |||  |||||||   ||   |||   |||||  ||||  |||||
     7257 CTGAGTAGCTGAGACTACAAGCGCAAGCCACCATGCCTGGCTAATTTTTTTTTTTGAGAC 7198

      694 AGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGACC 753
          |||||||  |  ||||   |||||  |||||||  ||  |||   ||||||   |  |||||||  |||
     7197 GGAGTCTTACTCTGTTGCCCAGTCTGGAGTGCAATGGTGTGATCTTGGCTCACTGCAACC 7138

      754 TCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAG 813
          ||   ||||||  |||||||||  ||||||||||||||||  ||||   ||||||  |  |||||||
     7137 TCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGAATTACAG 7078

      814 AAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATG 873
          |    ||  |  |||    |||||||||||||||||||||||||  |  |  |||   |||
     7077 GCGCCTGCCACTACGCCTGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCATCATA 7018

      874 TTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAA 933
          ||||||  |||  |||  |||| ||  ||||||||||||||  ||  ||||||||||||||||
     7017 ATGGCCAGGCTGATCTCGAACTCCTGACCTCAGGTGATCTGCCTGCCTTGGCCTCCCAAA 6958

      934 GTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCT 976
          |||||||||||  ||||||||||||  ||  |||  |||||   |||
     6957 GTGCTGGGATTACAGGCGTGAGCCACCGTGCCTGGCCTCGGCT 6915


LT 13
9-949-002-584/c
quence 584, Application US/09949002
tent No. 6900016
NERAL INFORMATION:
PPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 584
;   LENGTH: 21308
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-002-584

  Query Match             16.9%;  Score 315.8;  DB 3;  Length 21308;
  Best Local Similarity   82.0%;  Pred. No. 5.9e-77;
  Matches  401;  Conservative    0;  Mismatches   82;  Indels    6;  Gaps    3;

Qy        505 TTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT-GTCGCCCAGGCTGGAG 563
              ||| ||||| ||||| ||||||||||||||||||| | ||||| || ||||||||||||||
Db       3157 TTTCTTTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTCGTTGCCCAGGCTGGAG 3098

Qy        564 TGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCT 623
              |||| |||||||||||| || |||  |||||||||||||| | | ||||| |||||||||
Db       3097 TGCAATGGCATGATCTCGGCTCACCACAACCTCTGCCTCTCAGGTTCAAACGATTCTCCT 3038

Qy        624 GCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---TTTTT 680
              ||||||||||||| |||||||||||||||||| |    || ||||||  ||| ||  |||||
Db       3037 GCCTCAGCCTCCCGAGTAGCTGGGATTACAAGCATGCGGCACCACGCCCGGCTAATTTTT 2978

Qy        681 TTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCA 740
              || |||| ||||| |||| |  | ||||||||||||||||||| ||||||||||||||||
Db       2977 TTTTTTTTGAGACGGAGTGTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCATGATCTCA 2918

Qy        741 GTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTA 800
              | ||||||| |||||||||||||||||||||  |||||||| ||||||  |||| |||||
Db       2917 GCTCACTGCAACCTCCACCTCCCGGGTTCAAGTGATTCTCCTACCTCAGCCTCCCGAGTA 2858

Qy        801 GCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATG 860
              ||| ||| ||||| ||||  || ||||| ||||||||||||||||||||| ||||||
Db       2857 GCTGGGACTACAGGAGTGTGCCACCACGCTCAGCTAATTTTTGTATTTTTAATAGAGACA 2798

Qy        861 CGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCC 920
              | |||  |||| ||||||| ||| |||| |||| || |||||||  ||||| ||   ||
Db       2797 GGGTTTCACCATATTGGCCAGGCTGGTCTTGAACTCCTGACCTCA--TGATCTGCCCACC 2740

Qy        921 TTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTC 980
              |  |||||||||||||||||||||||||||||||| ||||||||  | ||| || ||    |  |
Db       2739 TCTGCCTCCCAAAGTGCTGGGATTGCAGGTGTGAGCCACTGTGCCTGGTCTCTTTTTTTT 2680

Qy        981 CTTTAGTCT 989
              |||  | |
Db       2679 TTTTTTTTT 2671


RESULT 14
US-09-949-002-784/c
; Sequence 784, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
;  TITLE OF INVENTION:  AND USES THEREOF
;  FILE REFERENCE: CL000790
;  CURRENT APPLICATION NUMBER: US/09/949,002
;  CURRENT FILING DATE:  2000-01-28
;  PRIOR APPLICATION NUMBER: 60/231,401
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 10823
;  SOFTWARE: FastSEQ for Windows Version 4.0
```

) ID NO 784
LENGTH: 21407
YPE: DNA
RGANISM: Human
)-949-002-784

ry Match 16.9%; Score 315.8; DB 3; Length 21407;
t Local Similarity 82.0%; Pred. No. 5.9e-77;
ches 401; Conservative 0; Mismatches 82; Indels 6; Gaps 3;

```
Qy       505 TTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT-GTCGCCCAGGCTGGAG 563
Db      3254 TTTCTTTTTTTTTTTTTTTTTTTTGAGACAGAGTTTCGCTCTCGTTGCCCAGGCTGGAG 3195

Qy       564 TGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCT 623
Db      3194 TGCAATGGCATGATCTCGGCTCACCACAACCTCTGCCTCTCAGGTTCAAACGATTCTCCT 3135

Qy       624 GCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---TTTTT 680
Db      3134 GCCTCAGCCTCCCGAGTAGCTGGGATTACAAGCATGCGGCACCACGCCCGGCTAATTTTT 3075

Qy       681 TTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCA 740
Db      3074 TTTTTTTTGAGACGGAGTGTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCATGATCTCA 3015

Qy       741 GTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTA 800
Db      3014 GCTCACTGCAACCTCCACCTCCCGGGTTCAAGTGATTCTCCTACCTCAGCCTCCCGAGTA 2955

Qy       801 GCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATG 860
Db      2954 GCTGGGACTACAGGAGTGTGCCACCACGCTCAGCTAATTTTTGTATTTTTAATAGAGACA 2895

Qy       861 CGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCC 920
Db      2894 GGGTTTCACCATATTGGCCAGGCTGGTCTTGAACTCCTGACCTCA--TGATCTGCCCACC 2837

Qy       921 TTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTC 980
Db      2836 TCTGCCTCCCAAAGTGCTGGGATTGCAGGTGTGAGCCACTGTGCCTGGTCTCTTTTTTTT 2777

Qy       981 CTTTAGTCT 989
Db      2776 TTTTTTTTT 2768
```

.T 15
)-949-016-11751
quence 11751, Application US/09949016
ent No. 6812339
ERAL INFORMATION:
PLICANT: VENTER, J. Craig et al.
TLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
LE REFERENCE: CL001307
RRENT APPLICATION NUMBER: US/09/949,016
RRENT FILING DATE: 2000-04-14
IOR APPLICATION NUMBER: 60/241,755
IOR FILING DATE: 2000-10-20
IOR APPLICATION NUMBER: 60/237,768
IOR FILING DATE: 2000-10-03
IOR APPLICATION NUMBER: 60/231,498
IOR FILING DATE: 2000-09-08
MBER OF SEQ ID NOS: 207012
FTWARE: FastSEQ for Windows Version 4.0
ID NO 11751
ENGTH: 40742
YPE: DNA
RGANISM: Human
)-949-016-11751

ry Match 16.7%; Score 312.8; DB 3; Length 40742;
Best Local Similarity 78.7%; Pred. No. 5.6e-76;
Matches 389; Conservative 0; Mismatches 97; Indels 8; Gaps 1;

```
Qy       502 TTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
Db     35103 TTTTTTCCTTTTTTTTTTTTTTTTAATTTGAGACAAGGTCTCGCTCTTTCACCCAGGTTGG 35162

Qy       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
Db     35163 AGTACAGTGGCGCGATCTTGGCTTACTGCAACCTTCGCCTCCCAGATTCAAGCGATTCTC 35222

Qy       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
Db     35223 ATGTCTCAGCCTCCCAAGTAGCTGGGATTATAGGCGTGCACCACCACACCAAGCTAATTT 35282

Qy       676 --TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCAT 733
Db     35283 TTTTTTTTTTTTTTGGTGACACAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGCAC 35342

Qy       734 GATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTC 793
Db     35343 GATTTTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGTGATTCTCCTGCCTTAGCCTC 35402

Qy       794 CTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGT 853
Db     35403 CCGAGTAGCTGGGACTACAGGCACATGCCACCGTGCCCAGCTAATTTTTGCATTTTTAGT 35462

Qy       854 AGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCC 913
Db     35463 AGAGACAGCATTTCACCATGTTGGCCAGGCAGGTCTTGAACTCCTGACCTCAGGTGATCC 35522

Qy       914 GCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGA 973
Db     35523 ACCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGAGCCTGGCCAAT 35582

Qy       974 GCTACTCCTTTAGT 987
Db     35583 TTTGTATTTTTAGT 35596
```

Search completed: June 22, 2006, 00:45:44
Job time : 382 secs

GenCore version 5.1.9

ucleic - nucleic search, using sw model

on: June 21, 2006, 22:25:04 ; Search time 2431 Seconds
(without alignments)
9452.027 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ence: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ing table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ched: 18892170 seqs, 6143817638 residues

l number of hits satisfying chosen parameters: 37784340

num DB seq length: 0
num DB seq length: 2000000000

processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : Published_Applications_NA_Main:*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

t o.	Score	% Query Match	Length	DB	ID	Description
1	1870	100.0	1870	10	US-10-502-279-26	Sequence 26, Appl
2	333.4	17.8	998	12	US-10-301-480-552165	Sequence 552165,
3	333.4	17.8	998	12	US-10-301-480-1165574	Sequence 1165574,
4	330.8	17.7	33362	9	US-10-388-838-36	Sequence 36, Appl
5	326.4	17.5	968	12	US-10-301-480-605733	Sequence 605733,
6	326.4	17.5	968	12	US-10-301-480-1219142	Sequence 1219142,
7	324.8	17.4	31116	6	US-10-087-192-1660	Sequence 1660, Ap
8	324.8	17.4	58822	8	US-10-052-482-46	Sequence 46, Appl
9	323.6	17.3	162450	6	US-10-071-179-1	Sequence 1, Appli
10	323.6	17.3	162450	7	US-10-126-704-1	Sequence 1, Appli
11	323.6	17.3	162450	13	US-11-132-838-1	Sequence 1, Appli
12	322.8	17.3	16963	8	US-10-741-601-5756	Sequence 5756, Ap
13	322.8	17.3	16963	9	US-10-741-600-17945	Sequence 17945, A
14	322.8	17.3	16963	10	US-10-995-561-13467	Sequence 13467, A
15	322.6	17.3	296405	6	US-10-087-192-1036	Sequence 1036, Ap
16	322.4	17.2	40000	10	US-10-995-561-13513	Sequence 13513, A
17	320.6	17.1	146656	15	US-11-121-086-68	Sequence 68, Appl
c 18	320.4	17.1	4957	7	US-10-074-024-753	Sequence 753, App
c 19	320.4	17.1	4961	7	US-10-074-024-752	Sequence 752, App
c 20	320.4	17.1	41907	3	US-09-967-013-5	Sequence 5, Appli
c 21	320.4	17.1	126001	7	US-10-175-492-13	Sequence 13, Appl
c 22	320.4	17.1	153142	15	US-11-121-086-27	Sequence 27, Appl
c 23	320.4	17.1	201309	16	US-11-114-798-51	Sequence 51, Appl
c 24	320.4	17.1	272022	11	US-10-330-773-102	Sequence 102, App
c 25	320.4	17.1	344805	9	US-10-779-271-1	Sequence 1, Appli
c 26	320.4	17.1	354592	10	US-10-737-082-70	Sequence 70, Appl
c 27	320.4	17.1	354592	10	US-10-765-790-70	Sequence 70, Appl
c 28	320.4	17.1	1080000	10	US-10-928-446A-1	Sequence 1, Appli
c 29	320.4	17.1	1080000	10	US-10-928-446A-181	Sequence 181, App
c 30	320.4	17.1	1080000	10	US-10-928-446A-183	Sequence 183, App
c 31	320.4	17.1	1080000	10	US-10-928-446A-185	Sequence 185, App
c 32	320.4	17.1	1080000	10	US-10-928-446A-187	Sequence 187, App
c 33	320.4	17.1	1080000	10	US-10-928-446A-189	Sequence 189, App
c 34	320.4	17.1	1080000	10	US-10-928-446A-191	Sequence 191, App
c 35	320.4	17.1	1080000	10	US-10-928-446A-193	Sequence 193, App
c 36	320.4	17.1	1080000	10	US-10-928-446A-195	Sequence 195, App
c 37	320.4	17.1	1080000	10	US-10-928-446A-197	Sequence 197, App
c 38	320.4	17.1	1080000	10	US-10-928-446A-199	Sequence 199, App
c 39	320.4	17.1	1080000	10	US-10-928-446A-201	Sequence 201, App
c 40	320.2	17.1	216929	8	US-10-741-601-5727	Sequence 5727, Ap
c 41	319.8	17.1	6093	7	US-10-074-024-535	Sequence 535, App
42	318	17.0	96649	3	US-09-956-712-10	Sequence 10, Appl
43	318	17.0	96649	8	US-10-633-913-10	Sequence 10, Appl
c 44	317.4	17.0	120096	15	US-11-121-086-24	Sequence 24, Appl
45	316	16.9	34875	9	US-10-775-169-316	Sequence 316, App

ALIGNMENTS

RESULT 1
US-10-502-279-26
; Sequence 26, Application US/10502279
; Publication No. US20050084840A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; APPLICANT: Hideki ENDOH
; APPLICANT: Ryosuke NAKANO
; APPLICANT: Eiji KUROSAKI
; APPLICANT: Miyuki KATO
; APPLICANT: Hiroyuki YOKOTA
; APPLICANT: Kazunori INABE
; TITLE OF INVENTION: METHOD FOR SCREENING A DRUG AMELIORATING INSULIN RESISTANCE
; FILE REFERENCE: Q82704
; CURRENT APPLICATION NUMBER: US/10/502,279
; CURRENT FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: JP 2002-013721
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: JP 2002-257703
; PRIOR FILING DATE: 2002-09-03
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1870)
; OTHER INFORMATION:
US-10-502-279-26

Query Match 100.0%; Score 1870; DB 10; Length 1870;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1870; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCCTGAAATAACTGCTTCT 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 ATCTGTGTCCTAGAAAGTACCCACGCAGACAATCTACAGGGTCCTGAAATAACTGCTTCT 60

```
 61 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 120
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 AGTTCTAGATTTCACCTCATTGTCCTGTGGGTAGTCCCTCCTTATTGCCCAGCTCCACTG 120

121 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 180
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 GCAAAGTTGGCCTCTCCTGAGCCTCTTGGATGAACCTGATTTCCATGTCCTCATGGGCCA 180

181 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 240
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 GGATATGGACTAGGAAAGTTACATCCAGAGAGAAGGCAGGGCTTGGTAAAATCTGAGCAG 240

241 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 300
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 CATTCTTGCCTATTCTTCCTGACCAGGGAGAATAACTAGGCAGAGCAATTCATATGCAAG 300

301 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 360
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 AAAACTTCTAGGTCAAGAGTACTTGGAAGAAGTGACCAGAAATGAACAGAAAAACATTAC 360

361 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 420
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 CTGGGAAAATATTCAAGGAGGAAGAAAAAAATATGCAGGGTCAGTGGCTTTCTTCACCTT 420

421 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 480
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 GGGTCTCAATCAGCCCTTCTAAGAGACCTGATATTCTGCTTGCTCCTGAGGCCAGAGCTG 480

481 CAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTTGAGACAGAGTC 540
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 CAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTTGAGACAGAGTC 540

541 TTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCC 600
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 TTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCC 600

601 TCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCAC 660
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 TCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCAC 660

661 GCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 GCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720

721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780

781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840

841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900

901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960

961 GCGCCAGGCCTGAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACG 1020
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
961 GCGCCAGGCCTGAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACG 1020

1021 CTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTAC 1080
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 CTTTACATGCCATCTCTCCCCTAGTCCCAAGGTCTTCTCTGGACTGGTTCCTTAATTTAC 1080

1081 CATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAGA 1140
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 CATCTCTCAATCAGAAGCGCCTTTACAAAGGATATCAGGATAGTATTTTTTGGTCAAAGA 1140

1141 TAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATAA 1200
```

```
     ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TAACTCTTCCCCCAGAACTCAGTAGGGTCTCTGATGGCATCAGTAAGGGCAATTTCATAA 1200

Qy   1201 AGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 AGCAACCAGGTTCCTGCCTGTTTACAGGCAGCTCCAGTTGTTGGGAAGTGAGCAAGCAAG 1260

Qy   1261 TGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGGA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TGAGGGGAAGCCAGACTTTCTGGAGCTTCCAGCCCTAGTTCTGCCACCCAGGCTTCGGGA 1320

Qy   1321 GGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGGT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGGCCCCACAAAGTCTGAAATCTACGTGCCCCACCCCCCAAAGCCCCTTCCAGCCCAGGT 1380

Qy   1381 TCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAGT 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TCTGGTTTAGCACTAGGCTCAAGAACCTCAGGCCTGAATTCTACTTCCCGTGCATTAAGT 1440

Qy   1441 CCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTGC 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 CCCACCTTAACATCCCATTCTTATGAAGTCATCCTGTCCTGCGGGGCAATTCTGTGTTGC 1500

Qy   1501 TGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 TGCTGGGGCCTGGTCACTGATAGAGAAGGGCGTGGGCTGGGCCTACCCGGCAGGAGCCGC 1560

Qy   1561 CTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAATC 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CTCACCCCATCAGGCCCAGATCCTCTAAGACAAGACAGGGATAAAGTCTCCCGTGGAATC 1620

Qy   1621 CCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCCA 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 CCCTCCCGCCTCATATTCCCTGACCCTGCCCTCCCTGCGAGAACTCCAGTCCGGGCTCCA 1680

Qy   1681 CAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACCT 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CAAAGGGCGCGGTAAGAAACGTACACTTCCGAGGGGCAGACAAGGAGGGGGCGTTCACCT 1740

Qy   1741 ACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ACACCCCGGCCGGGCTATGGTCGGGAACAGGGTGAGGGGCCCAAGCCCCGGCTGCGGCAG 1800

Qy   1801 GCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GCGGGCGCCCTTACCCATGCTGACTCTGGAGCGGGCCGGGTGGCGCTGCCCATCCCCGGA 1860

Qy   1861 TCCTTCGTAG 1870
          ||||||||||
Db   1861 TCCTTCGTAG 1870
```

RESULT 2
US-10-301-480-552165/c
; Sequence 552165, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 552165
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapien

0-301-480-552165

ery Match 17.8%; Score 333.4; DB 12; Length 998;
t Local Similarity 81.6%; Pred. No. 2.2e-89;
ches 399; Conservative 0; Mismatches 86; Indels 4; Gaps 1;

```
Qy      502 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
Db      498 TTTATTTATTTATTTATTTATTTATTTTGAGACAGAGTCTCACTCTGTCTCCCAGGCTGG 439

Qy      562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
Db      438 AGTTCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCTCAGGTTCAAGTGATTCTT 379

Qy      622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----TT 677
Db      378 CTGCCTCAGCCTCCTGAGTAGCTGCAATTACAGGCGTGCACCACCACACCTGGCTAATTT 319

Qy      678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
Db      318 TTTTTTTTTTTGAGATGGAGTCTCGCTCTGTTGCCCAGGTTGGAGTGCAGTGGCGTGATC 259

Qy      738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
Db      258 TCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCTA 199

Qy      798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
Db      198 GTAGCTGGGACTACAGGCGTGCGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAG 139

Qy      858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
Db      138 ATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCC 79

Qy      918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTA 977
Db       78 ACCTCAGCCTCCCAAAGTGCTGGGATTTCAGCCGTGAGCCACCGTGCCCGGCCTATTTTT 19

Qy      978 CTCCTTTAG 986
Db       18 ATTTTTTAG 10
```

LT 3
0-301-480-1165574/c
quence 1165574, Application US/10301480
blication No. US20060057564A1
NERAL INFORMATION:
PPLICANT: Wang, David G.
TLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
TLE OF INVENTION: in the Human Genome
LE REFERENCE: 108827.137
URRENT APPLICATION NUMBER: US/10/301,480
URRENT FILING DATE: 2002-11-21
RIOR APPLICATION NUMBER: US 10/215,598
RIOR FILING DATE: 2002-08-09
RIOR APPLICATION NUMBER: US 60/311,695
RIOR FILING DATE: 2001-08-10
UMBER OF SEQ ID NOS: 1226818
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 1165574
LENGTH: 998
TYPE: DNA
ORGANISM: Homo sapien
0-301-480-1165574

ery Match 17.8%; Score 333.4; DB 12; Length 998;
t Local Similarity 81.6%; Pred. No. 2.2e-89;
ches 399; Conservative 0; Mismatches 86; Indels 4; Gaps 1;

```
Qy      502 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
Db      498 TTTATTTATTTATTTATTTATTTATTTTGAGACAGAGTCTCACTCTGTCTCCCAGGCTGG 439

Qy      562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
Db      438 AGTTCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCTCAGGTTCAAGTGATTCTT 379

Qy      622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----TT 677
Db      378 CTGCCTCAGCCTCCTGAGTAGCTGCAATTACAGGCGTGCACCACCACACCTGGCTAATTT 319

Qy      678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
Db      318 TTTTTTTTTTTGAGATGGAGTCTCGCTCTGTTGCCCAGGTTGGAGTGCAGTGGCGTGATC 259

Qy      738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
Db      258 TCGGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCTA 199

Qy      798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
Db      198 GTAGCTGGGACTACAGGCGTGCGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAG 139

Qy      858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
Db      138 ATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCC 79

Qy      918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTA 977
Db       78 ACCTCAGCCTCCCAAAGTGCTGGGATTTCAGCCGTGAGCCACCGTGCCCGGCCTATTTTT 19

Qy      978 CTCCTTTAG 986
Db       18 ATTTTTTAG 10
```

RESULT 4
US-10-388-838-36/c
; Sequence 36, Application US/10388838
; Publication No. US20040180344A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/388,838
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 33362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-388-838-36

Query Match 17.7%; Score 330.8; DB 9; Length 33362;
Best Local Similarity 83.0%; Pred. No. 8.1e-88;
Matches 377; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```
Qy      517 TTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGA 576
Db    12186 TCTCTTTTTTTTTAAGACAGAGTCCTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGA 12127

Qy      577 TCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCC 636
Db    12126 TCTCGGCTCACTGCAATCTCCACCTCCTGGGTTCAAGCTATTCTCCTGCCTCAGCCTCCC 12067

Qy      637 AAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATTTTGGAGACAGA 696
Db    12066 GAGTAGCTGGGATTACAGGCATGCACCACCACGCTTGGCTGATTTTTATTTTGAGACGGA 12007

Qy      697 GTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCC 756
Db    12006 GTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGCGTGATCTTTGCTCACTACAACCTCC 11947
```

```
     757 ACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAG 816
         || ||||| |||||||| ||||| ||||||||| |||| |||||||| |||||||||
   11946 GCCGCCCGGATTCAAGCGATTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCA 11887

     817 TGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTG 876
         ||  || ||| |  | ||||||||||||||||||||||||||||| | |||  ||||||||
   11886 TGTGCCACCATGCCCGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTG 11827

     877 GCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTG 936
         || | ||| |||| |||| || |||||||||||| | |   ||| ||||||||||||||
   11826 GCAAGGCTGGTCTTGAACTCCTGACCTCAGGTGATACCCCCACCTCGGCCTCCCAAAGTG 11767

     937 CTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
         |||||||| |||| |||||||| ||||| |||
   11766 CTGGGATTACAGGTGTGAGCCACTGCGCCCAGCC 11733
```

```
;T 5
)-301-480-605733
;uence 605733, Application US/10301480
)lication No. US20060057564A1
)ERAL INFORMATION:
)PLICANT: Wang, David G.
:TLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
:TLE OF INVENTION:  in the Human Genome
:LE REFERENCE: 108827.137
)RRENT APPLICATION NUMBER: US/10/301,480
)RRENT FILING DATE:  2002-11-21
(IOR APPLICATION NUMBER: US 10/215,598
(IOR FILING DATE: 2002-08-09
(IOR APPLICATION NUMBER: US 60/311,695
(IOR FILING DATE: 2001-08-10
)MBER OF SEQ ID NOS: 1226818
)FTWARE: FastSEQ for Windows Version 4.0
) ID NO 605733
;ENGTH: 968
:YPE: DNA
)RGANISM: Homo sapien
)-301-480-605733

:ry Match            17.5%;  Score 326.4;  DB 12;  Length 968;
:t Local Similarity   83.5%;  Pred. No. 2.8e-87;
:ches  396;  Conservative    0;  Mismatches   71;  Indels    7;  Gaps    2;

      503 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 562
          ||  ||||||| ||||| |||||||||||||||||||  || ||| ||||||||||||
      473 TTGATTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCACTTTGTTGCCCAGGCTGGA 532

      563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
          || |||||||||||||  || |||| |||||||||||||| | || |||||||||
      533 GTACAGTGGCATGATCTGAGCTCACTACAACCTCTGCCTCCCAGGTTGAAGCGATTCAAG 592

      623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TT 677
          |||||||||||||||||||||||||||||| ||||||||||||||||||||    ||      ||
      593 TGCCTCAGCCTCCCAAGTAGCTGGGATTTCAGGTGCACGCCACCACACTTGGCTATTTTT 652

      678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
          ||||| |||| |||||||||||| || |||||||| |||||||||| |||||||   |||
      653 TTTTTTTTTTTGAGACAGAGTCTAGCTCTGTCACCGAGGCTGGAGTGCAGTGGCGCCATC 712

      738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
          || ||||| |||  |||||| |||||||||||||  |||||||||||||| |||||||||
      713 TCGGTTCAGTGCAGCCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCATTCTCCTGA 772

      798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
          ||||||  |||||||||     |||| ||| |  ||||||||||||||||||||||||
      773 GTAGCTGGGATTACAGGCACCCACCACCAAGCCCAGCTAATTTTTGTATTTTTAGTAGAG 832

      858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
          ||| | |||  |  | ||||| | ||| || | |||| || |||||||  ||||| ||
      833 ATGGGGTTTCACTGTATTGGCTAGGCTGGTTTCGAACTCCTGACCTCA--TGATCTGCCT 890

      918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
```

```
          ||||||||||||||||||||||| ||| |||||||||||||| |||||| |||||
Db        891 GCCTTGGCCTCCCAAAGTGCTGGAATTACAGGCGTGAGCCACCGCGCCTGGCCT 944
```

```
RESULT 6
US-10-301-480-1219142
; Sequence 1219142, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identifiction and Mapping of Single Nucleotide Polymorphisms
;  TITLE OF INVENTION:  in the Human Genome
;  FILE REFERENCE: 108827.137
;  CURRENT APPLICATION NUMBER: US/10/301,480
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 10/215,598
;  PRIOR FILING DATE: 2002-08-09
;  PRIOR APPLICATION NUMBER: US 60/311,695
;  PRIOR FILING DATE: 2001-08-10
;  NUMBER OF SEQ ID NOS: 1226818
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1219142
;   LENGTH: 968
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-10-301-480-1219142

  Query Match            17.5%;  Score 326.4;  DB 12;  Length 968;
  Best Local Similarity   83.5%;  Pred. No. 2.8e-87;
  Matches  396;  Conservative    0;  Mismatches   71;  Indels    7;  Gaps    2;

Qy        503 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 562
              ||  ||||||| ||||| |||||||||||||||||||  || ||| ||||||||||||
Db        473 TTGATTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCACTTTGTTGCCCAGGCTGGA 532

Qy        563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
              || |||||||||||||  || |||| |||||||||||||| | || |||||||||
Db        533 GTACAGTGGCATGATCTGAGCTCACTACAACCTCTGCCTCCCAGGTTGAAGCGATTCAAG 592

Qy        623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TT 677
              |||||||||||||||||||||||||||||| ||||||||||||||||||||    ||      ||
Db        593 TGCCTCAGCCTCCCAAGTAGCTGGGATTTCAGGTGCACGCCACCACACTTGGCTATTTTT 652

Qy        678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
              ||||| |||| |||||||||||| || |||||||| |||||||||| |||||||   |||
Db        653 TTTTTTTTTTTGAGACAGAGTCTAGCTCTGTCACCGAGGCTGGAGTGCAGTGGCGCCATC 712

Qy        738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
              || ||||| |||  |||||| |||||||||||||  |||||||||||||| |||||||||
Db        713 TCGGTTCAGTGCAGCCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCATTCTCCTGA 772

Qy        798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
              ||||||  |||||||||     |||| ||| |  ||||||||||||||||||||||||
Db        773 GTAGCTGGGATTACAGGCACCCACCACCAAGCCCAGCTAATTTTTGTATTTTTAGTAGAG 832

Qy        858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
              ||| | |||  |  | ||||| | ||| || | |||| || |||||||  ||||| ||
Db        833 ATGGGGTTTCACTGTATTGGCTAGGCTGGTTTCGAACTCCTGACCTCA--TGATCTGCCT 890

Qy        918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
              ||||||||||||||||||||||| ||| |||||||||||||| |||||| |||||
Db        891 GCCTTGGCCTCCCAAAGTGCTGGAATTACAGGCGTGAGCCACCGCGCCTGGCCT 944
```

```
RESULT 7
US-10-087-192-1660
; Sequence 1660, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
```

```
LE REFERENCE: 529452000122
URRENT APPLICATION NUMBER: US/10/087,192
URRENT FILING DATE:  2002-03-01
RIOR APPLICATION NUMBER: US 09/747,377
RIOR FILING DATE: 2000-12-22
RIOR APPLICATION NUMBER: US 09/798,586
RIOR FILING DATE: 2001-03-02
UMBER OF SEQ ID NOS: 2059
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 1660
ENGTH: 31116
YPE: DNA
RGANISM: Homo sapiens
0-087-192-1660

ery Match             17.4%;  Score 324.8;  DB 6;  Length 31116;
st Local Similarity   77.5%;  Pred. No. 5.1e-86;
tches  409;  Conservative    0;  Mismatches  112;  Indels    7;  Gaps    1;

       502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
           ||||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
     13037 TTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 13096

       562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
           | ||||||||  ||||||  || ||| ||||||||||||||||| | |||||| |||||||
     13097 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 13156

       622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
           ||||||||||||||  |||||||||||| |||||||||  | ||||  || |||  ||
     13157 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 13216

       676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
            || |||| |||| ||||  ||| ||  | ||||  ||||||||||||| || |||||||
     13217 TTTCTTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 13276

       735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
           ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||||| ||||
     13277 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 13336

       795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
           |||||||| |||||||||  |  | || |||||| | |||||||||||| ||||||||||
     13337 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 13396

       855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
           ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
     13397 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 13456

       915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
           |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
     13457 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 13516

       975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
           | ||   ||  | | |   |||| |   |     |  | | | |||
     13517 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 13564


LT 8
0-052-482-46/c
quence 46, Application US/10052482
blication No. US20040072264A1
NERAL INFORMATION:
PPLICANT: Engelhard, Eric
PPLICANT:  Morris, David
TLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
LE REFERENCE: A-71087/RMS/DCF
URRENT APPLICATION NUMBER: US/10/052,482
URRENT FILING DATE:  2002-08-15
RIOR APPLICATION NUMBER: US 09/747,377
RIOR FILING DATE: 2000-12-22
RIOR APPLICATION NUMBER: US 09/798,586
RIOR FILING DATE: 2001-03-02
UMBER OF SEQ ID NOS: 241
OFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 46
;   LENGTH: 58822
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (4679)..(5145)
;   OTHER INFORMATION: "n" at positions 4679 to 5145 can be any base
US-10-052-482-46

  Query Match             17.4%;  Score 324.8;  DB 8;  Length 58822;
  Best Local Similarity   77.5%;  Pred. No. 7.1e-86;
  Matches  409;  Conservative    0;  Mismatches  112;  Indels    7;  Gaps    1;

Qy        502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
              ||||||||||||| ||| |  | ||||||||||||||||||  |||| || |||||| |||
Db      58246 TTTTTTTTTTTTTTTTGTTATGTTTTTTGAGACAGAGTCTCACTCTATCCCCCAGGATGG 58187

Qy        562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
              | ||||||||  ||||||  || ||| ||||||||||||||||| | |||||| |||||||
Db      58186 AATGCAGTGGTGTGATCTTGGCTCACAGCAACCTCTGCCTCCCAGGTTCAAGTGATTCTC 58127

Qy        622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
              ||||||||||||||  |||||||||||| |||||||||  | ||||  || |||  ||
Db      58126 CTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCCCACCACTACGCCCGGCTAATTT 58067

Qy        676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
               || |||| |||| ||||  ||| ||  | ||||  ||||||||||||| || |||||||
Db      58066 TTTCTTTTTTTTTTGAGATGGAGCCTCACTCTGTTGCCCAGGCTGGAGTGCAATGGCATG 58007

Qy        735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
              ||||||| ||||||| |||||||||| || ||||||||  |||||||||||||||| ||||
Db      58006 ATCTCAGCTCACTGCAACCTCCACCTTCCAGGTTCAAGGGATTCTCCTGCCTCAGCCTCC 57947

Qy        795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
              |||||||| |||||||||  |  | || |||||| | |||||||||||| ||||||||||
Db      57946 CGAGTAGCTGGGATTACAGGTGCTCGCCACCACGTCCGGCTAATTTTTGCATTTTTAGTA 57887

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              ||||   | |||  ||| |||||||| ||| |||| |||| || ||||||||||||||
Db      57886 GAGACAGGGTTTCACCACGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCTA 57827

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
              |  | |||||| ||| |||||||||||||| |||| |||||||| |||||| |||||
Db      57826 CCTGTCTTGGCTTCCTAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCCGGCCTAGC 57767

Qy        975 CTACTCCTTTAGTCTCTGGAAAGACTGCGGCTCAGAGAAATCAACGCT 1022
              | ||   ||  | | |   |||| |   |     |  | | | |||
Db      57766 TTGCTTTCTTTTTTTTTTCTGAGACAGTCTCCTTCTGTCACCCAGGCT 57719


RESULT 9
US-10-071-179-1/c
; Sequence 1, Application US/10071179
; Publication No. US20030108882A1
; GENERAL INFORMATION:
;  APPLICANT: Bougueleret, Lydie
;  TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
;  TITLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
;  FILE REFERENCE: GENSET.031A
;  CURRENT APPLICATION NUMBER: US/10/071,179
;  CURRENT FILING DATE:  2002-02-07
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/345,882
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-06-30
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/091,315
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/111,909
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 140
;  SOFTWARE: Patent.pm
; SEQ ID NO 1
;   LENGTH: 162450
;   TYPE: DNA
```

```
)RGANISM: Homo sapiens
'EATURE:
)AME/KEY: allele
.OCATION: 72794
)THER INFORMATION: 5-124-273 : polymorphic base A or G
'EATURE:
:AME/KEY: allele
.OCATION: 88073
:THER INFORMATION: 5-127-261 : polymorphic base A or C
'EATURE:
:AME/KEY: allele
.OCATION: 90842
:THER INFORMATION: 99-1437-325 : polymorphic base A or G
'EATURE:
:AME/KEY: allele
.OCATION: 93714
'THER INFORMATION: 5-128-60 : polymorphic base deletion of GT
'EATURE:
:AME/KEY: allele
.OCATION: 97122
)THER INFORMATION: 99-1442-224 : polymorphic base G or T
'EATURE:
)AME/KEY: allele
.OCATION: 97152
)THER INFORMATION: 5-129-144 : polymorphic base deletion of T
'EATURE:
)AME/KEY: allele
.OCATION: 99098
)THER INFORMATION: 5-130-257 : polymorphic base A or G
'EATURE:
)AME/KEY: allele
.OCATION: 99117
)THER INFORMATION: 5-130-276 : polymorphic base A or G
'EATURE:
)AME/KEY: allele
.OCATION: 103806
)THER INFORMATION: 5-131-395 : polymorphic base A or T
'EATURE:
)AME/KEY: allele
.OCATION: 106940
:THER INFORMATION: 5-133-375 : polymorphic base insertion of A
'EATURE:
:AME/KEY: allele
.OCATION: 108106
:THER INFORMATION: 5-135-155 : polymorphic base insertion of A
'EATURE:
:AME/KEY: allele
.OCATION: 108149
:THER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
'EATURE:
:AME/KEY: allele
.OCATION: 108308
:THER INFORMATION: 5-135-357 : polymorphic base A or G
'EATURE:
:AME/KEY: allele
.OCATION: 108471
)THER INFORMATION: 5-136-174 : polymorphic base C or T
'EATURE:
)AME/KEY: allele
.OCATION: 134134
)THER INFORMATION: 5-140-120 : polymorphic base C or T
'EATURE:
)AME/KEY: allele
.OCATION: 134362
)THER INFORMATION: 5-140-348 : polymorphic base insertion of A
'EATURE:
)AME/KEY: allele
.OCATION: 134374
)THER INFORMATION: 5-140-361 : polymorphic base insertion of CA
'EATURE:
)AME/KEY: allele
.OCATION: 146328
)THER INFORMATION: 5-143-84 : polymorphic base A or G
```

```
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97130..97177
; OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97130..97177
; OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99075..99121
; OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99075..99121
; OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99094..99140
; OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
; FEATURE:
```

```
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177

ery Match            17.3%;  Score 323.6;  DB 6;  Length 162450;
t Local Similarity   81.2%;  Pred. No. 2.8e-85;
tches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

      488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
          | ||  || |  |   || ||||||| ||||   || ||||||||||||||||||||||
   102008 TAAGGAGCAGTAGTAATTCTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

      548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
          |||||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
   101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

      608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
          ||||||| || || |||||||||||||| |||||||||||| |||||  |  |||||||
   101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

      668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
          | |||||||  ||||||||||||||  |||  |||||| || |||||||||||||||||| |
   101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

      726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
          |||||| ||||||| | ||||||| | |||| |||||||||| |  | ||||||||||
   101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709

      786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
          |||| |||||||||||| ||| |||||  || |||| ||| |  | |||||||||  |
   101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

      844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
          | |||||||||||||   | |||  ||||||| |||| | | |||| || | || |||||
   101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

      904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
          ||||||||| ||  |||||||||||||||||||||||||| |||||||||||| ||||
   101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

      964 CCAGGCCTGAGCTA 977
          ||  |||  | | |
```

```
Db     101528 CCCAGCCACACCCA 101515

RESULT 10
US-10-126-704-1/c
; Sequence 1, Application US/10126704
; Publication No. US20030170647A1
; GENERAL INFORMATION:
;  APPLICANT: Bougueleret, Lydie
;  TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
;  TITLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
;  FILE REFERENCE: 44.US5.DIV
;  CURRENT APPLICATION NUMBER: US/10/126,704
;  CURRENT FILING DATE:  2002-04-20
;  PRIOR APPLICATION NUMBER: US 60/091,315
;  PRIOR FILING DATE: 1998-06-30
;  PRIOR APPLICATION NUMBER: US 60/111,909
;  PRIOR FILING DATE: 1998-12-10
;  NUMBER OF SEQ ID NOS: 140
;  SOFTWARE: Patent.pm
; SEQ ID NO 1
;   LENGTH: 162450
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 72794
;   OTHER INFORMATION: 5-124-273  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 88073
;   OTHER INFORMATION: 5-127-261  :  polymorphic  base  A  or  C
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 90842
;   OTHER INFORMATION: 99-1437-325  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 93714
;   OTHER INFORMATION: 5-128-60  :  polymorphic  base  deletion  of  GT
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97122
;   OTHER INFORMATION: 99-1442-224  :  polymorphic  base  G  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97152
;   OTHER INFORMATION: 5-129-144  :  polymorphic  base  deletion  of  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99098
;   OTHER INFORMATION: 5-130-257  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99117
;   OTHER INFORMATION: 5-130-276  :  polymorphic  base  A  or  G
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103806
;   OTHER INFORMATION: 5-131-395  :  polymorphic  base  A  or  T
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106940
;   OTHER INFORMATION: 5-133-375  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108106
;   OTHER INFORMATION: 5-135-155  :  polymorphic  base  insertion  of  A
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108149
;   OTHER INFORMATION: 5-135-198  :  polymorphic  base  insertion  of  GTTT
;   FEATURE:
```

```
)AME/KEY: allele
)OCATION: 108308
)THER INFORMATION: 5-135-357  :  polymorphic  base  A  or  G
'EATURE:
)AME/KEY: allele
)OCATION: 108471
)THER INFORMATION: 5-136-174  :  polymorphic  base  C  or  T
'EATURE:
)AME/KEY: allele
)OCATION: 134134
)THER INFORMATION: 5-140-120  :  polymorphic  base  C  or  T
'EATURE:
)AME/KEY: allele
)OCATION: 134362
)THER INFORMATION: 5-140-348  :  polymorphic  base  insertion  of  A
'EATURE:
)AME/KEY: allele
)OCATION: 134374
)THER INFORMATION: 5-140-361  :  polymorphic  base  insertion  of  CA
'EATURE:
)AME/KEY: allele
)OCATION: 146328
)THER INFORMATION: 5-143-84  :  polymorphic  base  A  or  G
'EATURE:
)AME/KEY: allele
)OCATION: 146345
)THER INFORMATION: 5-143-101  :  polymorphic  base  A  or  C
'EATURE:
)AME/KEY: allele
)OCATION: 150329
)THER INFORMATION: 5-145-24  :  polymorphic  base  A  or  G
'EATURE:
)AME/KEY: allele
)OCATION: 160031
)THER INFORMATION: 5-148-352  :  polymorphic  base  G  or  T
'EATURE:
)AME/KEY: allele
)OCATION: 72771..72817
)THER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
'EATURE:
)AME/KEY: allele
)OCATION: 72771..72817
)THER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
'EATURE:
)AME/KEY: allele
)OCATION: 88050..88096
)THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
'EATURE:
)AME/KEY: allele
)OCATION: 88050..88096
)THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
'EATURE:
)AME/KEY: allele
)OCATION: 90819..90865
)THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
'EATURE:
)AME/KEY: allele
)OCATION: 90819..90865
)THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
'EATURE:
)AME/KEY: allele
)OCATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
'EATURE:
)AME/KEY: allele
)OCATION: 93690..93736
)THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
'EATURE:
)AME/KEY: allele
)OCATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
'EATURE:
)AME/KEY: allele
;   LOCATION: 97099..97145
;   OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97130..97177
;   OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 97130..97177
;   OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99075..99121
;   OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99075..99121
;   OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99094..99140
;   OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 99094..99140
;   OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103783..103828
;   OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 103783..103828
;   OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106918..106966
;   OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 106918..106966
;   OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108127..108177
;   OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108127..108177
;   OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
;   FEATURE:

  Query Match            17.3%;  Score 323.6;  DB 7;  Length 162450;
  Best Local Similarity  81.2%;  Pred. No. 2.8e-85;
  Matches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              | ||  || | |   || ||||||| |||| || ||||||||||||||||||||||||
Db     102008 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
              |||||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
Db     101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

Qy        608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
```

```
          ||||||| || || |||||||||||||| |||||||||||| |||||  |  |||||||
   101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

      668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
          | |||||||  ||||||  ||||  |||  |||||| || ||||||||||||||||||| |
   101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

      726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
          |||||| ||||||| | ||||||| | |||| |||||||||| |   | ||||||||||
   101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709

      786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
          |||| |||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
   101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

      844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
          | |||||||||||||   | |||  ||||||| |||| | | |||| || | || |||||
   101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

      904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
          ||||||||| ||  |||||||||||||||||||||||||||| |||||||||||||| ||||
   101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

      964 CCAGGCCTGAGCTA 977
          ||  |||  | | |
   101528 CCCAGCCACACCCA 101515
```

```
LT 11
:-132-838-1/c
quence 1, Application US/11132838
blication No. US20050221371A1
NERAL INFORMATION:
PPLICANT: Bougueleret, Lydie
TLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING
TLE OF INVENTION:  PROTEIN (RBP-7)
TLE OF INVENTION:  AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
LE REFERENCE: GENSET.031A
URRENT APPLICATION NUMBER: US/11/132,838
URRENT FILING DATE:  2005-05-19
RIOR APPLICATION NUMBER: US/09/345,882
RIOR FILING DATE: 1999-06-30
RIOR APPLICATION NUMBER: US 60/091,315
RIOR FILING DATE: 1998-06-30
RIOR APPLICATION NUMBER: US 60/111,909
RIOR FILING DATE: 1998-12-10
UMBER OF SEQ ID NOS: 140
OFTWARE: Patent.pm
Q ID NO 1
LENGTH: 162450
YPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273  :  polymorphic  base  A  or  G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261  :  polymorphic  base  A  or  C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325  :  polymorphic  base  A  or  G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60  :  polymorphic  base  deletion  of  GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224  :  polymorphic  base  G  or  T
FEATURE:
```

```
;  NAME/KEY: allele
;  LOCATION: 97152
;  OTHER INFORMATION: 5-129-144  :  polymorphic  base  deletion  of  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 99098
;  OTHER INFORMATION: 5-130-257  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 99117
;  OTHER INFORMATION: 5-130-276  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 103806
;  OTHER INFORMATION: 5-131-395  :  polymorphic  base  A  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 106940
;  OTHER INFORMATION: 5-133-375  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108106
;  OTHER INFORMATION: 5-135-155  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108149
;  OTHER INFORMATION: 5-135-198  :  polymorphic  base  insertion  of  GTTT
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108308
;  OTHER INFORMATION: 5-135-357  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 108471
;  OTHER INFORMATION: 5-136-174  :  polymorphic  base  C  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134134
;  OTHER INFORMATION: 5-140-120  :  polymorphic  base  C  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134362
;  OTHER INFORMATION: 5-140-348  :  polymorphic  base  insertion  of  A
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 134374
;  OTHER INFORMATION: 5-140-361  :  polymorphic  base  insertion  of  CA
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 146328
;  OTHER INFORMATION: 5-143-84  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 146345
;  OTHER INFORMATION: 5-143-101  :  polymorphic  base  A  or  C
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 150329
;  OTHER INFORMATION: 5-145-24  :  polymorphic  base  A  or  G
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 160031
;  OTHER INFORMATION: 5-148-352  :  polymorphic  base  G  or  T
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 72771..72817
;  OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: 72771..72817
;  OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
;  FEATURE:
;  NAME/KEY: allele
```

```
.OCATION: 88050..88096
:THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
'EATURE:
!AME/KEY: allele
.OCATION: 88050..88096
:THER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
'EATURE:
!AME/KEY: allele
.OCATION: 90819..90865
:THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
'EATURE:
!AME/KEY: allele
.OCATION: 90819..90865
:THER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
'EATURE:
!AME/KEY: allele
.OCATION: 93690..93736
:THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
'EATURE:
!AME/KEY: allele
.OCATION: 93690..93736
:THER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
'EATURE:
!AME/KEY: allele
.OCATION: 97099..97145
)THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
'EATURE:
!AME/KEY: allele
.OCATION: 97099..97145
:THER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
'EATURE:
!AME/KEY: allele
.OCATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
'EATURE:
!AME/KEY: allele
.OCATION: 97130..97177
)THER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
'EATURE:
!AME/KEY: allele
.OCATION: 99075..99121
:THER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
'EATURE:
:AME/KEY: allele
.OCATION: 99075..99121
:THER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
'EATURE:
:AME/KEY: allele
.OCATION: 99094..99140
:THER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
'EATURE:
:AME/KEY: allele
.OCATION: 99094..99140
'THER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
'EATURE:
:AME/KEY: allele
.OCATION: 103783..103828
:THER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
'EATURE:
!AME/KEY: allele
,OCATION: 103783..103828
)THER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
'EATURE:
!AME/KEY: allele
,OCATION: 106918..106966
)THER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
'EATURE:
!AME/KEY: allele
,OCATION: 106918..106966
:THER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
'EATURE:
!AME/KEY: allele
.OCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108084..108130
;   OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 108127..108177
;   OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
;   FEATURE:
;   NAME/KEY: allele

  Query Match             17.3%;  Score 323.6;  DB 13;  Length 162450;
  Best Local Similarity   81.2%;  Pred. No. 2.8e-85;
  Matches  401;  Conservative    0;  Mismatches   89;  Indels    4;  Gaps    2;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              | ||  || | |   || ||||||| ||||   || ||||||||||||||||||||||||
Db     102008 TAAGGAGCAGTAGTAATTCTTTTTTTGTTTTGTCTTGTTTTTGAGACAGAGTCTTGCTCT 101949

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
              |||||||||||||||||||||||||  | ||||| || |||||||||||| || ||| ||
Db     101948 GTCGCCCAGGCTGGAGTGCAGTGGTGTCATCTCGGCTCACTGCAACCTCCGCTTCCTGGG 101889

Qy        608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
              ||||||| || || |||||||||||||||| |||||||||||| |||||  | |||||||
Db     101888 TTCAAGCAATCCTGCTGCCTCAGCCTCCCTAGTAGCTGGGATAACAGGCACCTGCCACCA 101829

Qy        668 CACCCAGCT--TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTAC 725
              | |||||||  ||||||  ||||  |||  |||||| || |||||||||||||||||| |
Db     101828 CGCCCAGCTAATTTTTTTTTTTTTAGATGGAGTCTGGCTCTGTCACCCAGGCTGGAGTGC 101769

Qy        726 AGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCC 785
              |||||| ||||||| | ||||||| | |||| |||||||||||  |  | ||||||||||
Db     101768 AGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTGACACCATTCTCCTGCC 101709

Qy        786 TCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTT--G 843
              |||| |||||||||||| ||| |||||  || |||| ||| |  | ||||||||||  |
Db     101708 TCAGCCTCCTGAGTAGCTGGGACTACAGGCGTCCACCACCATGCCCGGCTAATTTTTTGG 101649

Qy        844 TATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCT 903
              | |||||||||||||   | |||  ||||||| |||| | | |||| || | || |||||
Db     101648 TGTTTTTAGTAGAGACAGGGTTTCACCATGTTAGCCAGGATGGTCTCGATCTCCTGACCT 101589

Qy        904 CAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCG 963
              ||||||||| ||  |||||||||||||||||||||||||||||| |||||||||||| ||||
Db     101588 CAGGTGATCTGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG 101529

Qy        964 CCAGGCCTGAGCTA 977
              ||  |||  | | |
Db     101528 CCCAGCCACACCCA 101515


RESULT 12
US-10-741-601-5756/c
; Sequence 5756, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  STENOSIS, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001500
;  CURRENT APPLICATION NUMBER: US/10/741,601
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 26415
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5756
;   LENGTH: 16963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-741-601-5756

  Query Match             17.3%;  Score 322.8;  DB 8;  Length 16963;
```

```
st Local Similarity   82.8%;  Pred. No. 1.5e-85;
tches  394;  Conservative    2;  Mismatches   69;  Indels   11;  Gaps    2;

      506 TTTTTTTTCTTTTTCTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
          ||| |||| ||||| |||||||| ||||  |||||||||||||| ||||| |||||||||
     7767 TTTGTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCTGTAGCCCAAGCTGGAGTG 7708

      566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
          ||||||||||||||| ||| |||||||| |  |||||||| |||||||  |||||||||
     7707 CAGTGGCATGATCTCGGCCTACTGCAACCACCACCTCCCGGGTTCAAGCAGTTCTCCTGC 7648

      626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATT 685
          |||||||||||||||||||||||||||||||  | ||||||||||||||| |||||  ||
     7647 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGCACCACCACACCCAGCTATTTTTGTT 7588

      686 TTG---------GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          |||         ||||| |||||||||  ||||| |||||||||||| |||||| |||||
     7587 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 7528

      737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
          ||| | |||||| |||||| ||||||||||||||  |||||||||||||| |||| |
     7527 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 7468

      797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
          |||||||  ||: |||||  | |  |||||||      |||||  ||||||||||||||||
     7467 AGTAGCTGGGMCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 7408

      855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
          |||| |   |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
     7407 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 7348

      915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          |  |||| |||||||||||||||||||||||||||||:||||||| |  ||| ||||
     7347 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCRTGAGCCACCCTGCCCGGCC 7292


LT 13
0-741-600-17945/c
quence 17945, Application US/10741600
blication No. US20050026169A1
NERAL INFORMATION:
PPLICANT: CARGILL, Michele et al.
TLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
LE REFERENCE: CL001499
URRENT APPLICATION NUMBER: US/10/741,600
URRENT FILING DATE:  2003-12-22
UMBER OF SEQ ID NOS: 73997
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 17945
ENGTH: 16963
YPE: DNA
ORGANISM: Homo sapiens
0-741-600-17945

ery Match             17.3%;  Score 322.8;  DB 9;  Length 16963;
st Local Similarity   82.8%;  Pred. No. 1.5e-85;
tches  394;  Conservative    2;  Mismatches   69;  Indels   11;  Gaps    2;

      506 TTTTTTTTCTTTTTCTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
          ||| |||| ||||| |||||||| ||||  |||||||||||||| ||||| |||||||||
     7767 TTTGTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCTGTAGCCCAAGCTGGAGTG 7708

      566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
          ||||||||||||||| ||| |||||||| |  |||||||| |||||||  |||||||||
     7707 CAGTGGCATGATCTCGGCCTACTGCAACCACCACCTCCCGGGTTCAAGCAGTTCTCCTGC 7648

      626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATT 685
          |||||||||||||||||||||||||||||||  | ||||||||||||||| |||||  ||
     7647 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGCACCACCACACCCAGCTATTTTTGTT 7588

      686 TTG---------GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
          |||         ||||| |||||||||  ||||| |||||||||||| |||||| |||||
     7587 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 7528

      737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
          ||| | |||||| |||||| ||||||||||||||  |||||||||||||| |||| |
     7527 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 7468

      797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
          |||||||  ||: |||||  | |  |||||||      |||||  ||||||||||||||||
     7467 AGTAGCTGGGMCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 7408

      855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
          |||| |   |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
     7407 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 7348

      915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          |  |||| |||||||||||||||||||||||||||||:||||||| |  ||| ||||
     7347 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCRTGAGCCACCCTGCCCGGCC 7292
```

```
Db       7587 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 7528

Qy        737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
              ||| | |||||| |||||| |||||||||||||||  |||||||||||||| |||| |
Db       7527 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 7468

Qy        797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
              |||||||  ||: |||||  | |  |||||||      |||||  ||||||||||||||||
Db       7467 AGTAGCTGGGMCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 7408

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              |||| |   |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
Db       7407 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 7348

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
              |  |||| |||||||||||||||||||||||||||||:||||||| |  ||| ||||
Db       7347 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCRTGAGCCACCCTGCCCGGCC 7292


RESULT 14
US-10-995-561-13467/c
; Sequence 13467, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;  TITLE OF INVENTION:  DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001559
;  CURRENT APPLICATION NUMBER: US/10/995,561
;  CURRENT FILING DATE:  2004-11-24
;  NUMBER OF SEQ ID NOS: 85702
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13467
;   LENGTH: 16963
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-995-561-13467

  Query Match             17.3%;  Score 322.8;  DB 10;  Length 16963;
  Best Local Similarity   82.8%;  Pred. No. 1.5e-85;
  Matches  394;  Conservative    2;  Mismatches   69;  Indels   11;  Gaps    2;

Qy        506 TTTTTTTTCTTTTTCTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
              ||| |||| ||||| |||||||| ||||  |||||||||||||| ||||| |||||||||
Db       7767 TTTGTTTTTTTTTTTTTTTTTTTTGGAGATGGAGTCTTGCTCTGTAGCCCAAGCTGGAGTG 7708

Qy        566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
              ||||||||||||||| ||| |||||||| |  |||||||| |||||||  |||||||||
Db       7707 CAGTGGCATGATCTCGGCCTACTGCAACCACCACCTCCCGGGTTCAAGCAGTTCTCCTGC 7648

Qy        626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTTTATT 685
              |||||||||||||||||||||||||||||||  | ||||||||||||||| |||||  ||
Db       7647 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGTGCACCACCACACCCAGCTATTTTTGTT 7588

Qy        686 TTG---------GAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGAT 736
              |||         ||||| |||||||||  ||||| |||||||||||| |||||| |||||
Db       7587 TTGTTTTGTTTTGAGACGGAGTCTTGCTGTGTCAGCCAGGCTGGAGTGCAGTGGTATGAT 7528

Qy        737 CTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTG 796
              ||| | |||||| |||||| |||||||||||||||  |||||||||||||| |||| |
Db       7527 CTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCG 7468

Qy        797 AGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTAGTA 854
              |||||||  ||: |||||  | |  |||||||      |||||  ||||||||||||||||
Db       7467 AGTAGCTGGGMCTACAGGCGCGTGCCTCCACACCTGGCTAATTTTTTTGTATTTTTAGTA 7408

Qy        855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
              |||| |   |||  |||||||||||| ||| || | |||| || ||||||| |||||| |
Db       7407 GAGACGAAGTTTCACCATGTTGGCCAGGCTGGTTTCGAACTCCTGACCTCAAGTGATCTG 7348

Qy        915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
              |  |||| |||||||||||||||||||||||||||||:||||||| |  ||| ||||
```

```
     7347 CCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCRTGAGCCACCCTGCCCGGCC 7292


;T 15
)-087-192-1036
juence 1036, Application US/10087192
)lication No. US20020182586A1
!ERAL INFORMATION:
'PLICANT: Morris, David W.
'PLICANT:  Engelhard, Eric K.
:TLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
:TLE OF INVENTION:  CANCER
:LE REFERENCE: 529452000122
)RRENT APPLICATION NUMBER: US/10/087,192
)RRENT FILING DATE:  2002-03-01
!IOR APPLICATION NUMBER: US 09/747,377
!IOR FILING DATE: 2000-12-22
!IOR APPLICATION NUMBER: US 09/798,586
!IOR FILING DATE: 2001-03-02
:MBER OF SEQ ID NOS: 2059
:FTWARE: FastSEQ for Windows Version 4.0
; ID NO 1036
.ENGTH: 296405
:YPE: DNA
:RGANISM: Homo sapiens
'EATURE:
:AME/KEY: misc_feature
.OCATION: (1)...(296405)
:THER INFORMATION: n = A,T,C or G
)-087-192-1036

'ry Match            17.3%;  Score 322.6;  DB 6;  Length 296405;
:t Local Similarity   79.7%;  Pred. No. 7.6e-85;
.ches  396;  Conservative    0;  Mismatches   94;  Indels    7;  Gaps    1;

      488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
          || | || |||    |   | |||||  |||  || |||||||| ||  ||||||  ||||
   117715 TGTGATGATGCAATATGAATCTTTTTTGTTTGTTTATTTTTTGGGATGGAGTCTCACTCT 117774

      548 GTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGA 607
          ||  |||||||||||||||||||| |  ||||| || | |||||||||||||||||| |
   117775 GTAACCCAGGCTGGAGTGCAGTGGTACAATCTCAGCTCGCTGCAACCTCTGCCTCCCCGG 117834

      608 TTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCA 667
          ||| ||| |||||||||||||||||||||||||||||||||||||||||  |  |||||||
   117835 TTCCAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCACCTGCCACCA 117894

      668 CACCCAGCT-------TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGG 720
            | |||||       ||| ||  ||| |||||||||||||| | ||||  ||||||||||
   117895 TGCTCAGCTAATTTTGTTTGTTTGTTTTGAGACAGAGTCTTACTCTGTGGCCCAGGCTGG 117954

      721 AGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTC 780
          ||| ||||||||  ||||| | || |||| |||||||||||||| ||||| ||| |||||
   117955 AGTGCAGTGGCACAATCTCGGCTCGCTGCAACCTCCACCTCCCCGGTTCCAGCTATTCTT 118014

      781 CTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTT 840
          |||||||||||||| ||||||| |||||| ||  |||  |  |||||  | ||||||||
   118015 CTGCCTCAGTCTCCCGAGTAGCTGGGATTATAGGCGTGTGCTACCACGCCCGGCTAATTT 118074

      841 TTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGA 900
          |||||||||||||||||| | | |||  |||||||||||| ||| |||| |||| || |
   118075 TTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGG 118134

      901 CCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATC 960
          ||||| ||||||| |  |||||||||||||||||||||||||| |||| |||||||| |
   118135 CCTCAAGTGATCCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACC 118194

      961 GCGCCAGGCCTGAGCTA 977
            ||| |||||||  ||
   118195 CTGCCCGGCCTGAAATA 118211


:h completed: June 21, 2006, 23:07:23
```

Job time : 2441 secs

GenCore version 5.1.9

:ucleic - nucleic search, using sw model

on: June 21, 2006, 22:25:40 ; Search time 59 Seconds
(without alignments)
7598.058 Million cell updates/sec

e: US-10-502-279-26
ect score: 1870
ence: 1 atctgtgtcctagaaagtac.....catccccggatccttcgtag 1870

ng table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ched: 296510 seqs, 119862409 residues

number of hits satisfying chosen parameters: 593020

num DB seq length: 0
num DB seq length: 2000000000

processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

base : Published_Applications_NA_New:*
1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lt o.	Score	% Query Match	Length	DB	ID	Description
1	296.6	15.9	4407	6	US-10-196-749-351	Sequence 351, App
2	284.8	15.2	12963	6	US-10-517-441-68	Sequence 68, Appl
3	284.4	15.2	1237	6	US-10-511-937-430	Sequence 430, App
4	284.4	15.2	2549	7	US-11-293-697-1413	Sequence 1413, Ap
5	282.4	15.1	3329	7	US-11-293-697-1502	Sequence 1502, Ap
6	271	14.5	7432	6	US-10-517-441-2	Sequence 2, Appli
7	270.8	14.5	2837	7	US-11-293-697-2282	Sequence 2282, Ap
8	270.8	14.5	394191	6	US-10-506-549-3	Sequence 3, Appli
9	269.6	14.4	118899	7	US-11-189-279-64	Sequence 64, Appl
10	269	14.4	10865	6	US-10-517-441-109	Sequence 109, App
11	266.6	14.3	7432	6	US-10-517-441-2	Sequence 2, Appli
12	262.2	14.0	3495	7	US-11-293-697-463	Sequence 463, App
13	260.8	13.9	2284	7	US-11-293-697-2393	Sequence 2393, Ap
14	257.8	13.8	3174	7	US-11-293-697-116	Sequence 116, App
15	257	13.7	2425	6	US-10-517-441-94	Sequence 94, Appl
16	256.8	13.7	2033	7	US-11-293-697-2350	Sequence 2350, Ap
17	254.4	13.6	176373	7	US-11-175-714-87	Sequence 87, Appl
18	253.2	13.5	11029	6	US-10-517-441-29	Sequence 29, Appl
19	247.4	13.2	6521	6	US-10-517-441-32	Sequence 32, Appl
20	242.4	13.0	152331	7	US-11-175-714-86	Sequence 86, Appl
21	239.8	12.8	1740	7	US-11-293-697-1833	Sequence 1833, Ap
22	239.2	12.8	2327	7	US-11-293-697-1071	Sequence 1071, Ap
23	238.8	12.8	2133	7	US-11-293-697-468	Sequence 468, App
24	237.6	12.7	2730	7	US-11-293-697-250	Sequence 250, App
25	236.4	12.6	2638	7	US-11-293-697-55	Sequence 55, Appl
c 26	232.6	12.4	16125	7	US-11-236-238-1	Sequence 1, Appli
c 27	231	12.4	2676	7	US-11-293-697-125	Sequence 125, App
28	230.6	12.3	2690	7	US-11-293-697-682	Sequence 682, App
29	228.6	12.2	3097	7	US-11-293-697-1206	Sequence 1206, Ap
c 30	228	12.2	56580	6	US-10-553-298-1	Sequence 1, Appli
c 31	225.6	12.1	2532	7	US-11-293-697-360	Sequence 360, App
32	222.8	11.9	1801	7	US-11-293-697-1829	Sequence 1829, Ap
c 33	222.8	11.9	118899	7	US-11-189-279-64	Sequence 64, Appl
c 34	222.4	11.9	1601	7	US-11-328-161-25	Sequence 25, Appl
c 35	222.4	11.9	1645	7	US-11-328-161-12	Sequence 12, Appl
c 36	221.4	11.8	1629	6	US-10-511-937-444	Sequence 444, App
37	220.6	11.8	56580	6	US-10-553-298-1	Sequence 1, Appli
38	220.4	11.8	2860	7	US-11-293-697-25	Sequence 25, Appl
c 39	220.2	11.8	6101	6	US-10-517-441-93	Sequence 93, Appl
c 40	220.2	11.8	27684	6	US-10-105-871-4	Sequence 4, Appli
41	220	11.8	2024	7	US-11-293-697-304	Sequence 304, App
42	217.6	11.6	2836	7	US-11-293-697-30	Sequence 30, Appl
43	217.2	11.6	11029	6	US-10-517-441-29	Sequence 29, Appl
c 44	217.2	11.6	128361	6	US-10-505-928-151	Sequence 151, App
c 45	216.6	11.6	2072	7	US-11-293-697-1005	Sequence 1005, Ap

ALIGNMENTS

RESULT 1
US-10-196-749-351
; Sequence 351, Application US/10196749
; Publication No. US20060094864A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C340
; CURRENT APPLICATION NUMBER: US/10/196,749
; CURRENT FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 351
; LENGTH: 4407
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-196-749-351

```
:ry Match            15.9%;  Score 296.6;  DB 6;  Length 4407;
:t Local Similarity   76.0%;  Pred. No. 2.3e-38;
:ches  379;  Conservative    0;  Mismatches  119;  Indels    1;  Gaps    1;

      473 CAGAGCTGCAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAG 532
          ||| |    || ||||    |  |  |  ||| ||| |||  || || ||||||||||||
     3503 CAGGGAGTGAGGGATTATCTTTTTTTTTTTTTCTTTCTTTCTTTCTTTTTTTTTTTTGAG 3562

      533 ACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAA 592
          ||||| ||| ||||||||||||||||||||||||||||| |||||  ||||| || |||||||
     3563 ACAGAATCTCGCTCTGTCGCCCAGGCTGGAGTGCAATGGCACAATCTCGGCTCACTGCAT 3622

      593 CCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTAC 652
          |||| |||||||||| |||||| ||||||| |||||||||||||  ||||||||||||||||
     3623 CCTCCGCCTCCCGGGTTCAAGTGATTCTCATGCCTCAGCCTCCTGAGTAGCTGGGATTAC 3682

      653 AGGTGCACGCCACCACACCCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACC 712
          |||  |  |||||||| |||||||  ||||  ||| |               |||  |
     3683 AGGCTCCTGCCACCACGCCCAGCTAATTTTTGTTTTGTTTTGTTTGGAGACAGAGTCTCG 3742

      713 CAGGCTGGAGTACAG-TGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAA 771
          |      |  |  | ||| ||||| |||| ||||||| |||| | || || |||||| |
     3743 CTATTGTCACCAGGGCTGGAATGATTTCAGCTCACTGCAACCTTCGCCACCTGGGTTCCA 3802

      772 GCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTC 831
          |||||||||||||||||| |||| ||||||||  ||||| ||       ||| |||||  |
     3803 GCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGAGATTATAGGCACCTACCACCACGCCC 3862

      832 AGCTAATTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGG 891
          |||||||||||||||||||||||||||| | |  |||  |||||||||||| ||| |||| |
     3863 GGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCG 3922

      892 AACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCG 951
          ||| || ||||| |||||||||  || |||||   ||||||||||||||||||| ||||||
     3923 AACTCCTGACCTTAGGTGATCCACTCGCCTTCATCTCCCAAAGTGCTGGGATTACAGGCG 3982

      952 TGAGCCATCGCGCCAGGCC 970
          ||||||| || ||| ||||
     3983 TGAGCCACCGTGCCTGGCC 4001


,T 2
)-517-441-68/c
quence 68, Application US/10517441
)lication No. US20060121467A1
IERAL INFORMATION:
)PLICANT: FOEKENS, John
)PLICANT:  HARBECK, Nadia
)PLICANT:  KOENIG, Thomas
)PLICANT:  MAIER, Sabine
)PLICANT:  MARTENS, John
)PLICANT:  MODEL, Fabian
)PLICANT:  NIMMRICH, Inko
)PLICANT:  RUJAN, Tamas
)PLICANT:  SCHMITT, Armin
)PLICANT:  SCHMITT, Manfred
)PLICANT:  LOOK, Maxime P.
)PLICANT:  MARX, Almuth
)PLICANT:  HOEFLER, Heinz
:TLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
:TLE OF INVENTION:  proliferative disorders
:LE REFERENCE: 47675-93
)RRENT APPLICATION NUMBER: US/10/517,441
)RRENT FILING DATE:  2004-12-11
(IOR APPLICATION NUMBER: PCT/EP2003/010881
(IOR FILING DATE: 2003-10-01
(IOR APPLICATION NUMBER: DE 10317955.0
(IOR FILING DATE: 2003-04-17
(IOR APPLICATION NUMBER: DE 10300096.8
(IOR FILING DATE: 2003-01-07
(IOR APPLICATION NUMBER: DE 10245779.4
(IOR FILING DATE: 2002-10-01
;  NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 68
;   LENGTH: 12963
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-517-441-68

  Query Match            15.2%;  Score 284.8;  DB 6;  Length 12963;
  Best Local Similarity   80.7%;  Pred. No. 1.4e-36;
  Matches  384;  Conservative    0;  Mismatches   82;  Indels   10;  Gaps    4;

Qy        502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
              | ||| |  ||| | ||| | |||||||||||  ||||||  |||||| ||||||||||
Db       6678 TATTTATCCTTTTTCTTTTTCTTTTTTTGAGATGGAGTCTCACTCTGTTGCCCAGGCTGG 6619

Qy        562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
              |||||| |||  ||||| | || || |||||| |||||||||  || ||||||||||||||
Db       6618 AGTGCAATGGTGTGATCCCAGCTCAGTGCAACTTCTGCCTCCTGGGTTCAAGCGATTCTC 6559

Qy        622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC----TT 677
              ||||||||| ||||| |||||||||||||| ||||||| | ||||||  ||||||     ||
Db       6558 CTGCCTCAGACTCCCGAGTAGCTGGGATTTCAGGTGCCCACCACCATGCCCAGCTAATTT 6499

Qy        678 TTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATC 737
              ||||| |||| ||||    |||||||| ||||| |||||||||||||| |||||||| ||||
Db       6498 TTTTTTTTTTTGAGATGAAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCACGATC 6439

Qy        738 TCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGA 797
              |  | ||||||| ||||  |||||||||||||||||||||  |||||||||  |||||||
Db       6438 TTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCAATT---CTGCCTCAGCCTCCTGA 6382

Qy        798 GTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAG 857
              |||||| |||||||||    | || ||||  | ||| ||||| ||||||||||||||||
Db       6381 GTAGCTGGGATTACAGGCAGGTGCCACCAC-ACCGGCTGATTTTCGTATTTTTAGTAGAG 6323

Qy        858 ATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTG 917
              |||   |||  |||  |||| || ||| |||| |||| || ||||||  |||||| ||
Db       6322 ATGGCGTTTCACCACATTGGTCAGGCTGGTCTTGAACTCCTGACCTC--GTGATCTGCCC 6265

Qy        918 GCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGA 973
              ||| |||||||||||||||||||||||||||||||||||||||| || |||  ||||||
Db       6264 ACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCAGCCTGA 6209


RESULT 3
US-10-511-937-430/c
; Sequence 430, Application US/10511937
; Publication No. US20060088836A1
; GENERAL INFORMATION:
;  APPLICANT: EXPRESSION DIAGNOSTICS, INC.
;  APPLICANT:  Wohlgemuth, Jay
;  APPLICANT:  Fry, Kirk
;  APPLICANT:  Woodward, Robert
;  APPLICANT:  Ly, Ngoc
;  APPLICANT:  Prentice, James
;  APPLICANT:  Morris, MacDonald
;  APPLICANT:  Rosenberg, Steven
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
;  TITLE OF INVENTION:  AND MONITORING TRANSPLANT REJECTION
;  FILE REFERENCE: 506612000104
;  CURRENT APPLICATION NUMBER: US/10/511,937
;  CURRENT FILING DATE:  2004-10-19
;  PRIOR APPLICATION NUMBER: PCT/US2003/012946
;  PRIOR FILING DATE: 2003-04-24
;  PRIOR APPLICATION NUMBER: US 10/131,831
;  PRIOR FILING DATE: 2002-04-24
;  PRIOR APPLICATION NUMBER: US 10/325,899
;  PRIOR FILING DATE: 2002-12-20
;  NUMBER OF SEQ ID NOS: 3117
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 430
;   LENGTH: 1237
;   TYPE: DNA
```

```
ORGANISM: Homo sapiens
.0-511-937-430

ery Match            15.2%;  Score 284.4;  DB 6;  Length 1237;
st Local Similarity  74.2%;  Pred. No. 2e-36;
tches  396;  Conservative    0;  Mismatches  116;  Indels   22;  Gaps    2;

       473 CAGAGCTGCAGAGATTGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTCTTTTTTTTTGAG 532
           ||| ||  ||||   |   ||      |   ||         ||| |  || || ||| |
      1018 CAGCGCCTCAGAAGCTCTTCTAGGCTTTAGTTGGGGAAGCTTTTTGTTGTTGTTGTTGTG 959

       533 ACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAA 592
           ||  ||||||   ||  |||| |||||||||||||| |||||||  |||||  || |||||||
       958 ACGGAGTCTCACTTTGTCACCCAGGCTGGAGTGAAGTGGCACAATCTCAGCTCACTGCAG 899

       593 CCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTAC 652
           ||||  |||||  ||  |||||  | ||| ||||||||  ||||||||  ||||||||||||||||
       898 CCTCCACCTCCTGGGTTCAAACAATTTTCCTGCCTTAGCCTCCCGAGTAGCTGGGATTAC 839

       653 AGGTGCACGCCACCACACCCAGC--------------------TTTTTTTATTTTGGAG 691
           |||  |   ||||||||  |||  ||                    |||||||  ||||  |||
       838 AGGCGTGGGCCACCACGCCCGGCTAATTTTTGTATTTTTTTTTTTTTTTTTTTTTTGAG 779

       692 ACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTCAGTTCACTGCGA 751
           ||  ||||||  ||  |||||  ||||||||||||||  |||||||   ||||||  | |||||||  |
       778 ACGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAA 719

       752 CCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTAC 811
            ||||  ||||||||||||||  ||  ||||||||||||||||  ||||  ||||||||  |||  |||
       718 GCTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTAC 659

       812 AGAAGTGCACCTCCACGTTCAGCTAA-TTTTTGTATTTTTAGTAGAGATGCGCTTTTGCC 870
           ||   |  | |  |||||   | |||||  ||||||||||||||||||||||||  | |||   |
       658 AGGCGCCCGCTACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCATC 599

       871 ATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCC 930
           ||||||||||  |||  ||||  ||||  ||  |||||||  |||||||  |    ||||||||||||
       598 ATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCCACCCACCTTGGCCTCCC 539

       931 AAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCCTTT 984
           ||||||||||||||  ||||||||||||||  |  ||  |  |||||  |  |  |||
       538 AAAGTGCTGGGATTACAGGCGTGAGCCACCACGTCCAGCCTGGGGAAGGTTTTT 485


LT 4
.1-293-697-1413
equence 1413, Application US/11293697
blication No. US20060105376A1
ENERAL INFORMATION:
PPLICANT: HELIX RESEARCH INSTITUTE
ITLE OF INVENTION: Novel full length cDNA
ILE REFERENCE: H1-A0106
URRENT APPLICATION NUMBER: US/11/293,697
URRENT FILING DATE:  2005-12-05
RIOR APPLICATION NUMBER: US/10/108,260
RIOR FILING DATE: 2002-03-28
UMBER OF SEQ ID NOS: 5458
OFTWARE: PatentIn Ver. 2.1
Q ID NO 1413
LENGTH: 2549
YPE: DNA
ORGANISM: Homo sapiens
.1-293-697-1413

ery Match            15.2%;  Score 284.4;  DB 7;  Length 2549;
st Local Similarity  78.3%;  Pred. No. 1.9e-36;
tches  382;  Conservative    0;  Mismatches   96;  Indels   10;  Gaps    3;

       509 TTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAG 568
           ||||||||||| |||  |||||| ||    |||||  ||||||||  ||||||||  ||||||
      1807 TTTTTCTTTTTTTTTCTTTTTGGGATTAAGTCTCGCTCTGTCACCCAGGCTAAAGTGCAT 1866

       569 TGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTC 628

           | |   ||||||  || ||||||||||||||||||||  || ||  |||  |||  |||  ||||||
       Db 1867 TTGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTTAAGTGATCCTCATGCCTC 1926

       Qy  629 AGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT--------TTTT 680
               |||||||||||||||||||||||||||||| |    |||||||   ||  |||         | ||
       Db 1927 AGCCTCCCAAGTAGCTGGGATTACAGGCGTGTGCCACCATGCCTGGCTAATTTTTGTATT 1986

       Qy  681 TTATTTTGGAGACAGAGTCTTGC-CCTGTCACCCAGGCTGGAGTACAGTGGCATGATCTC 739
               || |||| ||||  |||| |  | | |||  | ||||||||||| |||||||  |||||||
       Db 1987 TTTTTTTTGAGATGGAGTTTCACTCTTGTTGCTCAGGCTGGAGTGCAGTGGCGCGATCTC 2046

       Qy  740 AGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGT 799
                 | |||||||  ||||||  ||||||  |||||||||  ||||||||||||||||  ||||    ||
       Db 2047 GGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCTGGT 2106

       Qy  800 AGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTAGAGAT 859
               ||||  |||||||||    ||  ||  |||  |   |||||||  |||  ||||||  ||||||||||
       Db 2107 AGCTGGGATTACAGGCATGTGCCACCATGCCCAGCTAA-TTTCGTATTTCTAGTAGAGAT 2165

       Qy  860 GCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGC 919
               | | |||  ||||||  || ||  |||  |||||||||  ||  ||||||||||||||  ||  ||
       Db 2166 GGGGTTTCTCCATGTCGGTCAGGCTGGTCTGGAACTCCTGACCTCAGGTGATCTGCCTGC 2225

       Qy  920 CTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACT 979
               ||||||||||  ||||||||||||  |||||  |||||||  ||  |||  |||||     |
       Db 2226 CTTGGCCTCCCGAAGTGCTGGGATTACAGGCATGAGCCACCGTGCCCGGCCTTTTTTGTA 2285

       Qy  980 CCTTTAGT 987
                 ||||||
       Db 2286 TTTTTAGT 2293


RESULT 5
US-11-293-697-1502/c
; Sequence 1502, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1502
;   LENGTH: 3329
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-1502

  Query Match            15.1%;  Score 282.4;  DB 7;  Length 3329;
  Best Local Similarity  77.8%;  Pred. No. 3.7e-36;
  Matches  381;  Conservative    0;  Mismatches  101;  Indels    8;  Gaps    3;

       Qy  502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
               ||  ||   |||| ||||  |||||||||||||||| ||||  |||||||  | ||  |||||
       Db 1584 TTCTTCCCTTTTTTTTTCTTTTTTTTTTGAGACAGGGTCTCACTCTGTCACTCATGCTGG 1525

       Qy  562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
               |||||||||| |||||||  || |||||||||||||||||||  ||  |||||  |||  ||
       Db 1524 AGTGCAGTGGCGTGATCTCAGCTCACTGCAACCTCTGCCTCCTGGCCTCAAGTGATCCTT 1465

       Qy  622 CTGCCTCAG---CCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC--- 675
                 |||||||    |  ||||||||||||| | ||||||   |  ||||||   |||||
       Db 1464 TCGCCTCAGTCCCTACCCAAGTAGCTGGAACTACAGGCACGAACCACCATGACCAGCTAA 1405

       Qy  676 -TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATG 734
                | || | |||| ||||||||||||||| |||||||  |||| ||||   || ||
       Db 1404 ATTATTATTTTTTTTGAGACAGAGTCTTGCTCTGTCACTCAGGATGGAGGGGAGCGGTGCC 1345

       Qy  735 ATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCC 794
```

```
      ||||| | ||| ||| |||||| ||| || |||||||||||||||||| |||||| ||||
  1344 ATCTCGGCTCATTGCAACCTCCGCCT-CCAGGTTCAAGCAATTCTCCTACCTCAGCCTCC 1286

   795 TGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGTA 854
       ||||||| ||| ||||| | |||| ||||     ||||||||||||||||||||||
  1285 AAAGTAGCTGGGACTACAGGTGCACACCACCACTCCTGGCTAATTTTTGTATTTTTAGTG 1226

   855 GAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCG 914
       |||||| | ||| ||||||||||||| ||| ||||  ||| ||  |||||| |||||||
  1225 GAGATGGGGTTTCGCCATGTTGGCCAGGCTGGTCTCAAACTCCTAACCTCAAGTGATCCA 1166

   915 CTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAG 974
       || |||||||||||||||||||||||||| ||| |||||||||  | ||| ||||
  1165 CTCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGACGTGAGCCACTGTGCCTGGCCACTC 1106

   975 CTACTCCTTT 984
       ||   | |||
  1105 CTGTACTTTT 1096
```

```
;T 6
)-517-441-2/c
;uence 2, Application US/10517441
)lication No. US20060121467A1
;ERAL INFORMATION:
)PLICANT: FOEKENS, John
)PLICANT:  HARBECK, Nadia
)PLICANT:  KOENIG, Thomas
)PLICANT:  MAIER, Sabine
)PLICANT:  MARTENS, John
)PLICANT:  MODEL, Fabian
)PLICANT:  NIMMRICH, Inko
)PLICANT:  RUJAN, Tamas
)PLICANT:  SCHMITT, Armin
)PLICANT:  SCHMITT, Manfred
)PLICANT:  LOOK, Maxime P.
)PLICANT:  MARX, Almuth
)PLICANT:  HOEFLER, Heinz
:TLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
:TLE OF INVENTION:  proliferative disorders
:LE REFERENCE: 47675-93
)RRENT APPLICATION NUMBER: US/10/517,441
)RRENT FILING DATE:  2004-12-11
(IOR APPLICATION NUMBER: PCT/EP2003/010881
(IOR FILING DATE: 2003-10-01
(IOR APPLICATION NUMBER: DE 10317955.0
(IOR FILING DATE: 2003-04-17
(IOR APPLICATION NUMBER: DE 10300096.8
(IOR FILING DATE: 2003-01-07
(IOR APPLICATION NUMBER: DE 10245779.4
(IOR FILING DATE: 2002-10-01
)MBER OF SEQ ID NOS: 2147
) ID NO 2
;ENGTH: 7432
:YPE: DNA
)RGANISM: Homo Sapiens
)-517-441-2

:ry Match            14.5%;  Score 271;  DB 6;  Length 7432;
;t Local Similarity   76.4%;  Pred. No. 1.9e-34;
:ches  375;  Conservative    0;  Mismatches  105;  Indels   11;  Gaps    3;
```

```
      492 CTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCG 551
          ||| | |    |||| || |  |||| || |||||||||  || |||  |||||| |
     7239 CTGATTTTAACATTTTAATTATTATTTTTTTCTTTTTTGAGATGGAATCTCACTCTGTTG 7180

      552 CCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCA 611
          ||||||||||||||||||    |||| || |||||||||| |  |||||| | ||||
     7179 CCCAGGCTGGAGTGCAGTGGTGCAATCTTGGCTCACTGCAACCCCCACCTCCCAGGTTCA 7120

      612 AGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACAC- 670
          ||| |||||||||||||||||||| |||||||||||||||||  | |||||||||| |
     7119 AGCAATTCTCCTGCCTCAGCCTCCCCAGTAGCTGGGATTACAGGCACCCGCCACCACCCG 7060
```

```
Qy        671 -------CCAGCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGT 723
                   | ||| |  |||| |||| |||||| || ||||||||||||||||||||
Db       7059 GCTGATTTTTGTTTTCGTATTTTTTGAGATGGAGTCTCGCTCTGTCACCCAGGCTGGAGT 7000

Qy        724 ACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTG 783
              ||||||| || |||| | ||||||| |||||  ||||| |||||| ||||||||| |||
Db       6999 GCAGTGGCGTGTTCTCGGCTCACTGCAACCTCTGCCTCCAGGGTTCTAGCAATTCTTCTG 6940

Qy        784 CCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAAT-TTTT 842
              |||||| ||||  || |||| | || |||        || ||| |  | |||||| ||||
Db       6939 CCTCAGCCTCCCAAGGAGCTGGAATCACAAGCACCTGCCACCATGCCCCGGCTAATATTTT 6880

Qy        843 GTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACC 902
              |||||||||||||||||| | ||| ||||| |||||| ||| |||| |||| | ||||
Db       6879 GTATTTTTAGTAGAGATGGGGTTTCACCATGCTGGCCAGGCTGGTCTCGAACTTCTGACC 6820

Qy        903 TCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGC 962
              ||  |||||||||   |||||||||||||||||||| ||||| ||||| |||||||  |
Db       6819 TC--GTGATCCGCCCACCTTGGCCTCCCAAAGTGCTAGGATTACAGGCTTGAGCCACTGT 6762

Qy        963 GCCAGGCCTGA 973
              |||   |||||
Db       6761 GCCCCACCTGA 6751
```

```
RESULT 7
US-11-293-697-2282
; Sequence 2282, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2282
;   LENGTH: 2837
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-2282

  Query Match            14.5%;  Score 270.8;  DB 7;  Length 2837;
  Best Local Similarity   76.3%;  Pred. No. 2.3e-34;
  Matches  389;  Conservative    0;  Mismatches  107;  Indels   14;  Gaps    4;
```

```
Qy        492 CTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCG 551
              || ||  |   |||    |||   |||| | |||||||||||||| ||||  |||||| |
Db       1620 CTCCTCTTCTCTTTACAATTTAACTTTTTTCTTTTTTTGAGACAGGGTCTCACTCTGTTG 1679

Qy        552 CCCAGGCTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCA 611
              ||||| |||||||||||||| ||||||   || ||||| | |||| |||||| || ||||
Db       1680 CCCAGCCTGGAGTGCAGTGGTATGATCATGGCTCACTGTAGCCTCAGCCTCCTGGGTTCA 1739

Qy        612 AGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACC 671
              | |||||||||| ||||||||   |||||||||| ||||||||| | ||| ||||||
Db       1740 GGTGATTCTCCTGCTTCAGCCTCATGAGTAGCTGGGACTACAGGTGCTCACCATCACACC 1799

Qy        672 CAGCT-----------TTTTTTATTTTGGAGACAGAGTCT-TGCCCTGTCACCCAGGCTG 719
              |||||           || ||  |||| ||| |||||| |    | |||  |||||||||
Db       1800 CAGCTAATTATTATTATTATTATTTTTTGAGGCAGAGTTTCATTCTTGTTGCCCAGGCTG 1859

Qy        720 GAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCT 779
              |||| || |||||||||||| | ||||||| |||||  | |||| |||||||||  |||||
Db       1860 GAGTGCAATGGCATGATCTCGGCTCACTGCAACCTCTGCTTCCCAGGTTCAAGCGATTCT 1919

Qy        780 CCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATT 839
              |||||||||| |||| | |||||| |||||||||   ||  || ||| |  | |||||
Db       1920 CCTGCCTCAGCCTCCCGGGTAGCTGGGATTACAGGCATGAGCCACCATGCCCGGCTAA-T 1978
```

```
      840 TTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGG 899
          |||||||||||||||||| | | |||  |||||||||  || ||| |||| ||||  |
     1979 TTTGTATTTTTAGTAGAGACG-GGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTTCCA 2037

      900 ACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCAT 959
          |||||||||||  | |  |||||||||||||||||||||||||||| ||||||||||||
     2038 ACCTCAGGTGATTCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 2097

      960 CGCGCCAGGCCTGAGCTACTCCTTTAGTCT 989
          | |||  |||||   |  |  |||| | |
     2098 TGAGCCCAGCCTGCTATTGTTTTTTAATTT 2127


LT 8
0-506-549-3/c
quence 3, Application US/10506549
blication No. US20060100417A1
NERAL INFORMATION:
PPLICANT: APPLERA CORPORATION
TLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
TLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
TLE OF INVENTION:  AND USES THEREOF
LE REFERENCE: CL001361-US
URRENT APPLICATION NUMBER: US/10/506,549
URRENT FILING DATE:  2004-09-03
RIOR APPLICATION NUMBER: 60/361,343
RIOR FILING DATE: 2002-03-05
UMBER OF SEQ ID NOS: 4
OFTWARE: FastSEQ for Windows Version 4.0
Q ID NO 3
ENGTH: 394191
YPE: DNA
ORGANISM: Homo sapiens
EATURE:
AME/KEY: misc_feature
OCATION: (1)...(394191)
OTHER INFORMATION: n = A,T,C or G
0-506-549-3

ery Match            14.5%;  Score 270.8;  DB 6;  Length 394191;
t Local Similarity   78.2%;  Pred. No. 1.4e-34;
tches  377;  Conservative    0;  Mismatches   97;  Indels    8;  Gaps    4;

       518 TTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCATGAT 577
           || | ||||||||||||||||||  ||||||| |||||||| ||||||||||    ||||
    331446 TTTTATTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTAGAGTGCAGTGATGTGAT 331387

       578 CTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCA 637
           ||  || |||||||| |||  |||||| | | ||||| ||| |||     ||||||||||||
    331386 CTTGGCTCACTGCAATCTCCACCTCCTGCACTCAAGTGATACTCTCTTCTCAGCCTCCCA 331327

       638 AGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC-----TTTTTTTATTTTGGAGA 692
           ||||||||||  ||||||    | ||||||  ||||||     |||||||| |||| |||
    331326 GGTAGCTGGGACCACAGGTATGCACCACCATGCCCAGCTATTTTTTTTTTTTTTTTGAGG 331267

       693 CAGAGTCTTGC-CCTGTCACCCAGGCTGG-AGTACAGTGGCATGATCTCAGTTCACTGCG 750
             |||| | || | |||  |||||||||| ||| ||||||     |||  | |||||||
    331266 TGGAGTTTCGCTCTTGTTGCCCAGGCTGGAAGTGCAGTGGTGCAGTCTTTGCTCACTGCA 331207

       751 ACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTA 810
           |||||  |||||| |||||| |||||||||||||||||||| ||||  |||||||| ||||||
    331206 ACCTCTGCCTCCTGGGTTCGAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTA 331147

       811 CAGAAGTGCACCTCCACGTTCAGCTAA-TTTTTGTATTTTTAGTAGAGATGCGCTTTTGC 869
           |||  |  |||| |||||     ||||| |||||||||||||||||||| | ||||||
    331146 CAGGCGCCCACCACCACGCCTGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTTGC 331087

       870 CATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCC 929
           |||  |||||| ||| |||| |||| || ||||||||||||||| ||  |||  |||||||
    331086 CATACTGGCCAGGCTGGTCTTGAACTCCAGACCTCAGGTGATCTGCCTGCCACGGCCTCC 331027

       930 CAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGAGCTACTCCTTTAGTCT 989
```

```
           |||||||||||||||| ||||| ||||||| || |||  |||  |||| | |||| || |
Db     331026 CAAAGTGCTGGGATTACAGGCTTGAGCCACCGTGCCCAGCCCCAGCTATTTCTTTTGTAT 330967

Qy        990 CT 991
              ||
Db     330966 TT 330965


RESULT 9
US-11-189-279-64
; Sequence 64, Application US/11189279
; Publication No. US20060115829A1
; GENERAL INFORMATION:
;  APPLICANT: MAO, LI
;  APPLICANT:  WANG, JIE
;  APPLICANT:  LUO, WANG
;  TITLE OF INVENTION: A METHOD OF TREATING CANCER
;  FILE REFERENCE: UTXC:875US
;  CURRENT APPLICATION NUMBER: US/11/189,279
;  CURRENT FILING DATE:  2005-07-26
;  PRIOR APPLICATION NUMBER: 60/598,554
;  PRIOR FILING DATE: 2004-08-03
;  NUMBER OF SEQ ID NOS: 69
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
;   LENGTH: 118899
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-189-279-64

  Query Match            14.4%;  Score 269.6;  DB 7;  Length 118899;
  Best Local Similarity  78.0%;  Pred. No. 2.4e-34;
  Matches  382;  Conservative    0;  Mismatches   89;  Indels   19;  Gaps    4;

Qy        510 TTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGT 569
              |||| ||||| |||||||||||||||||| ||| |||||||| ||||||||||||||||||
Db       5250 TTTTTTTTTTTTTTTTTTTTGAGACAGATTCTGGCTCTGTCACCCAGGCTGGAGTGCAGT 5309

Qy        570 GGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGCCTCA 629
              ||||  ||||| || |||||||  ||||| |||||| ||| || |||||||| ||||||
Db       5310 GGCACAATCTCGGCTCACTGCAATTTCTGCTTCCCGGGTTCTAGTGATTCTCCAGCCTCA 5369

Qy        630 GCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCT------------ 676
              ||||||| | |||||||||  | ||||||  | |||||| ||||||||
Db       5370 GCCTCCCGAATAGCTGGGAATGCAGGTGGCCACCACCAAACCCAGCTAACTTTTGTCTTT 5429

Qy        677 --TTTTTTATTTTGGAGACAGAGTCTTGC-CCTGTCACCCAGGCTGGAGTACAGTGGCAT 733
              |||||| ||||||| || |||| |||| | | |   |||||||||||| || ||||
Db       5430 TTTTTTTTTTTTTGGAAACGGAGTTTTGCTCTTTTTGTCCAGGCTGGAGTGCAATGGCGA 5489

Qy        734 GATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTC 793
              |||||| | ||||||| |||||| ||||||||||||||||  |||||||| ||||| |||
Db       5490 GATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGTCTCAGCCTC 5549

Qy        794 CTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAATTTTTGTATTTTTAGT 853
              |  ||||||| |||||||||   ||  || ||||   | |   | ||||||| |||||||
Db       5550 CCAAGTAGCTGGGATTACAGGCATGTGCCACCAC--CCCGGCTAATTTTGTACTTTTAGT 5607

Qy        854 AGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGATCC 913
              ||||| | | |||  |||||||||| || ||| |||| |||| ||  ||||||||||||
Db       5608 AGAGACG-GGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACTCCCCACCTCAGGTGATCC 5666

Qy        914 GCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCTGA 973
              ||  |||| |||||||||||||||||| || |||||| |||||| ||| || ||||  |
Db       5667 GCCTGCCTCGGCCTCCCAAAGTGCTGGGTTTACAGGCGCGAGCCACCGCACCTGGCCCAA 5726

Qy        974 GCTACTCCTT 983
                |   |||
Db       5727 CTTTGTCTT 5736


RESULT 10
US-10-517-441-109
```

```
¡uence 109, Application US/10517441
>lication No. US20060121467A1
;ERAL INFORMATION:
>PLICANT: FOEKENS, John
>PLICANT:  HARBECK, Nadia
>PLICANT:  KOENIG, Thomas
>PLICANT:  MAIER, Sabine
>PLICANT:  MARTENS, John
>PLICANT:  MODEL, Fabian
>PLICANT:  NIMMRICH, Inko
>PLICANT:  RUJAN, Tamas
>PLICANT:  SCHMITT, Armin
>PLICANT:  SCHMITT, Manfred
>PLICANT:  LOOK, Maxime P.
>PLICANT:  MARX, Almuth
>PLICANT:  HOEFLER, Heinz
:TLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
:TLE OF INVENTION:  proliferative disorders
:LE REFERENCE: 47675-93
:RRENT APPLICATION NUMBER: US/10/517,441
:RRENT FILING DATE:  2004-12-11
;IOR APPLICATION NUMBER: PCT/EP2003/010881
;IOR FILING DATE: 2003-10-01
;IOR APPLICATION NUMBER: DE 10317955.0
;IOR FILING DATE: 2003-04-17
;IOR APPLICATION NUMBER: DE 10300096.8
;IOR FILING DATE: 2003-01-07
;IOR APPLICATION NUMBER: DE 10245779.4
;IOR FILING DATE: 2002-10-01
]MBER OF SEQ ID NOS: 2147
; ID NO 109
.ENGTH: 10865
:YPE: DNA
>RGANISM: Homo Sapiens
)-517-441-109

:ry Match            14.4%;  Score 269;  DB 6;  Length 10865;
;t Local Similarity   75.4%;  Pred. No. 3.8e-34;
:ches  378;  Conservative    0;  Mismatches  110;  Indels   13;  Gaps    3;

      504 TTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTC-TGTCGCCCAGGCTGGA 562
          ||| ||  | | ||| ||| |||||||||  |||| |  ||| | | |||||||||| ||
     6664 TTTAATTAATTTATTTATTTATTTTTGAGATGGAGTTTCACTCTTATTGCCCAGGCTAGA 6723

      563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
          ||||  |||| ||||||| || ||||||||||||||||||||| | ||||||| ||||| |
     6724 GTGCCATGGCGTGATCTCGGCTCACTGCAACCTCTGCCTCCCCGGTTCAAGCAATTCTTC 6783

      623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------- 675
          |||||||||||| | ||||||||||||||||||     |||||||  ||  ||
     6784 TGCCTCAGCCTCTCGAGTAGCTGGGATTACAGGCATGTGCCACCAAGCCTGGCTAATTTT 6843

      676 TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCATGA 735
          ||||||| |||| |||||||||||||  | |||||| |||||||||||||| |||||||  |
     6844 TTTTTTTTTTTTTGAGACAGAGTCTCACTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCAA 6903

      736 TCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTCCT 795
          |||| | ||||||| | |||  |||||| |||||| || |||||||| | ||||| ||||
     6904 TCTCGGCTCACTGCAAGCTCTGCCTCCCAGGTTCACGCCATTCTCCCGACTCAGCCTCCC 6963

      796 GAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAAT-----TTTTGTATTTTT 850
          |||||||  ||| |||||  |    || |||| | | ||||||      ||||||||||||
     6964 AAGTAGCTGGGACTACAGGCGCCTGCCACCACATCCGGCTAATTTTGATTTTGTATTTTT 7023

      851 AGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGA 910
          |||||||| | | |||   || |||||| || ||| |||| |||| || ||||||||||||
     7024 AGTAGAGACGGGGTTTCACCGTGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGA 7083

      911 TCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCC 970
          || ||   | |  |||||||||||||||||||||| ||||| |||||||  | ||| ||||
     7084 TCTGCCCACTTCTGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGTGCCCGGCC 7143

      971 TGAGCTACTCCTTTAGTCTCT 991
```

```
                | | |   || | | |
Db       7144 CTATATTTTTTAATATTTTAT 7164


RESULT 11
US-10-517-441-2
; Sequence 2, Application US/10517441
; Publication No. US20060121467A1
; GENERAL INFORMATION:
;  APPLICANT: FOEKENS, John
;  APPLICANT:  HARBECK, Nadia
;  APPLICANT:  KOENIG, Thomas
;  APPLICANT:  MAIER, Sabine
;  APPLICANT:  MARTENS, John
;  APPLICANT:  MODEL, Fabian
;  APPLICANT:  NIMMRICH, Inko
;  APPLICANT:  RUJAN, Tamas
;  APPLICANT:  SCHMITT, Armin
;  APPLICANT:  SCHMITT, Manfred
;  APPLICANT:  LOOK, Maxime P.
;  APPLICANT:  MARX, Almuth
;  APPLICANT:  HOEFLER, Heinz
;  TITLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
;  TITLE OF INVENTION:  proliferative disorders
;  FILE REFERENCE: 47675-93
;  CURRENT APPLICATION NUMBER: US/10/517,441
;  CURRENT FILING DATE:  2004-12-11
;  PRIOR APPLICATION NUMBER: PCT/EP2003/010881
;  PRIOR FILING DATE: 2003-10-01
;  PRIOR APPLICATION NUMBER: DE 10317955.0
;  PRIOR FILING DATE: 2003-04-17
;  PRIOR APPLICATION NUMBER: DE 10300096.8
;  PRIOR FILING DATE: 2003-01-07
;  PRIOR APPLICATION NUMBER: DE 10245779.4
;  PRIOR FILING DATE: 2002-10-01
;  NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 2
;   LENGTH: 7432
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-517-441-2

  Query Match            14.3%;  Score 266.6;  DB 6;  Length 7432;
  Best Local Similarity  76.6%;  Pred. No. 9.2e-34;
  Matches  382;  Conservative    0;  Mismatches  104;  Indels   13;  Gaps    4;

Qy        488 TGAGCTGCTGCTGCTTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCT 547
              ||  |  |||    ||   | |   | ||||| |||||||||||||| ||||||  ||||
Db       1695 TGTACGTCTGTGTTTTGCATATGCATTTTTTTTTTTTTTTTTGAGACGGAGTCTCACTCT 1754

Qy        548 GTCGCCCAGGCTGGAGTGCAGTGGCAT---------GATCTCTGCCCACTGCAACCTCTG 598
              ||| |||||||||||||||||||| |         |||||  || ||||||| | || |
Db       1755 GTCACCCAGGCTGGAGTGCAGTGGTAGCAGTGGTGCGATCTTGGCTCACTGCATCATCCG 1814

Qy        599 CCTCCCGGATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGC 658
              ||| || | |||||| |||||||||||||||| || |  ||||  ||||| |||||   |
Db       1815 CCTACCCGTTTCAAGGGATTCTCCTGCCTCAGTCTTCAGAGTATTTGGGACTACAGACAC 1874

Qy        659 ACGCCACCACACCCAGCT-TTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGC 717
              |||||||||  ||  ||| |||||| |||| ||||| |||||| || |||| ||||||||
Db       1875 ACGCCACCATGCCTGGCTATTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTTACCCAGGC 1934

Qy        718 TGGAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATT 777
              |||||| ||||||| ||||||  | ||||||| || ||| |||||||||||||| || |||
Db       1935 TGGAGTGCAGTGGCGTGATCTTGGCTCACTGCAACGTCCGCCTCCCGGGTTCACGCCATT 1994

Qy        778 CTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA 837
              |||||||||||| |||| |||||||| ||| ||||| ||  |||| |||||     |||||
Db       1995 CTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGAGCCCACCACCACGCCTGGCTAA 2054

Qy        838 -TTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCC 896
              |||||||||||||| |||||| | | ||| ||| |||| |||| | | |||| |   |
Db       2055 TTTTTTGTATTTTTACTAGAGACGGGGTTTCGCCGTGTTAGCCAGGATGGTCTCCATATC 2114
```

```
 897 CGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGC 956
     | ||||||  |||||||||  |||| ||||| ||||||||||  |||||  ||||||||||
2115 CTGACCTC--GTGATCCGCCTGCCTCGGCCTTCCAAAGTGCTAGGATTATAGGCGTGAGC 2172

 957 CATCGCGCCAGGCCTGAGC 975
     ||  ||||| ||||   ||
2173 CACTGCGCCTGGCCAATGC 2191


LT 12
:.-293-697-463
quence 463, Application US/11293697
blication No. US20060105376A1
NERAL INFORMATION:
PPLICANT: HELIX RESEARCH INSTITUTE
:TLE OF INVENTION: Novel full length cDNA
:LE REFERENCE: H1-A0106
URRENT APPLICATION NUMBER: US/11/293,697
URRENT FILING DATE:  2005-12-05
RIOR APPLICATION NUMBER: US/10/108,260
RIOR FILING DATE: 2002-03-28
UMBER OF SEQ ID NOS: 5458
OFTWARE: PatentIn Ver. 2.1
Q ID NO 463
ENGTH: 3495
:YPE: DNA
ORGANISM: Homo sapiens
:.-293-697-463

ery Match             14.0%;  Score 262.2;  DB 7;  Length 3495;
at Local Similarity   75.6%;  Pred. No. 4.7e-33;
tches  377;  Conservative    0;  Mismatches   93;  Indels   29;  Gaps    3;

 502 TTTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGG 561
     ||| ||||| ||  ||||     ||||||||||||||||| ||||||| ||||||||||
 103 TTTATTTTTATTTATTTTTAAAATTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTGG 162

 562 AGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTC 621
     |||| |||| |   ||||| |  ||||||| |||   ||||| || ||||||| ||||||
 163 AGTGGAGTGCCGCAATCTCGGTTCACTGCAGCCTCCACCTCCTGGGTTCAAGCAATTCTC 222

 622 CTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC------ 675
      |||||||||| |||  || ||||||||||||||||||   |||||||| ||  ||
 223 ATGCCTCAGCTTCCTGAGAAGCTGGGATTACAGGTGTGTGCCACCACGCCAGGCTAATTT 282

 676 -------------------TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGG 716
                         ||| |||||||| ||||||||||| ||  |||| ||||||
 283 TTGTATTTTTATTTATTTATTTATTTATTTTTGAGACAGAGTCTCGCTTTGTCGCCCAGG 342

 717 CTGGAGTACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAAT 776
     |||| || |||||||| |||||| | ||||||| | |||| |||||| |||||| || ||
 343 CTGGTGTGCAGTGGCACGATCTCGGCTCACTGCAAGCTCCGCCTCCCAGGTTCACGCCAT 402

 777 TCTCCTGCCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTA 836
     ||| |||||||||| |||| |||||||| ||| |||||  |  ||   ||| |    ||||
 403 TCTTCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGCCCATTACCATGCCTGGCTA 462

 837 A--TTTTTGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAAC 894
     |  |||||||||||||||||||||||| | |||  |  |||| |||| | | |||| || |
 463 ATTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACTGTGTTAGCCAGGATGGTCTCGATC 522

 895 CCCGGACCTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGA 954
      || | |||   ||||||| || |||||||||||||||||||||||||||  |||| |||
 523 TCCTGCCCT--TGTGATCCACTCGCCTTGGCCTCCCAAAGTGCTGGGATTATAGGCTTGA 580

 955 GCCATCGCGCCAGGCCTGA 973
     |||| ||||||||||||  |
 581 GCCACCGCGCCAGGCCTAA 599


LT 13
:.-293-697-2393
```

```
; Sequence 2393, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2393
;   LENGTH: 2284
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-293-697-2393

  Query Match             13.9%;  Score 260.8;  DB 7;  Length 2284;
  Best Local Similarity   77.3%;  Pred. No. 8.1e-33;
  Matches  377;  Conservative    0;  Mismatches   87;  Indels   24;  Gaps    4;

Qy        503 TTTTTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGA 562
              ||||| ||| | || || ||||||| |||||||||||||| |||||||| |||||||| ||
Db       1431 TTTTTCTTTCTTTTCTTTTTTTTTTCTGAGACAGAGTCTCGCTCTGTCTCCCAGGCTAGA 1490

Qy        563 GTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCC 622
              |||||||||||||||| || |||||||||||||| ||||||| |||| ||  |||| |
Db       1491 GTGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGTCTCCCGGGTTCAGGCAGTTCTTC 1550

Qy        623 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGCTTTTTT- 681
              |||||||||||||||||||||||||||||||||||||| | | ||||||||||||  |||
Db       1551 TGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCCACTACCACACCCAGCTAATTTG 1610

Qy        682 ------------------TATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGT 723
                                | |||| |||| ||||| ||| |||||||| |||||||||
Db       1611 TGTGTGTGTGTGTGTGTTTTTTTGAGATGGAGTCCTGCTCTGTCACCCCGGCTGGAGT 1670

Qy        724 ACAGTGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTG 783
              ||||||  |||||||| ||||||| | |||| |||||||||||||  | |||||||||
Db       1671 GCAGTGGTGTGATCTCAGCTCACTGCAAGCTCCGCCTCCCGGGTTCACACTATTCTCCTG 1730

Qy        784 CCTCAGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCT--AATTTT 841
              |||||| |||| |||| ||| ||| ||||| |    || ||| |    |||    ||||
Db       1731 CCTCAGCCTCCCGAGTGGCTGGGACTACAGGCGCCTGCCACCATGCCTGGCTGGTTTTTT 1790

Qy        842 TGTATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGAC 901
              ||| |||| |||||||||    | |||  || ||||||||| | | || | || | ||| |||
Db       1791 TGTGTTTTAAGTAGAGACAGGGTTTCACCGTGTTGGCCAGGATGGT-TCGATCTCCTGAC 1849

Qy        902 CTCAGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCG 961
              ||||  |||||  |    |||| || ||||||  |||||| |||| |||||||||||| ||
Db       1850 CTCA--TGATCTGGCCGCCTCGGTCTCCCAGGGTGCTGAGATTACAGGCGTGAGCCAGCG 1907

Qy        962 CGCCAGGC 969
              ||| |||
Db       1908 TGCCTGGC 1915


RESULT 14
US-11-293-697-116
; Sequence 116, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: Novel full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/11/293,697
;  CURRENT FILING DATE:  2005-12-05
;  PRIOR APPLICATION NUMBER: US/10/108,260
;  PRIOR FILING DATE: 2002-03-28
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
```

; ID NO 116
.ENGTH: 3174
:YPE: DNA
:RGANISM: Homo sapiens
.-293-697-116

:ry Match 13.8%; Score 257.8; DB 7; Length 3174;
:t Local Similarity 72.8%; Pred. No. 2.3e-32;
:ches 375; Conservative 0; Mismatches 132; Indels 8; Gaps 3;

```
      498 CTGCTTTTTTTTTTTTCTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGG 557
          ||   || || ||||| |||||   |||||||||||||  ||||  |||||  |||||||
     1043 CTTTCTTCTTCTTTTTTTTTTTAAATTTTTTTGAGACATGGTCTCACTCTGCTGCCCAGG 1102

      558 CTGGAGTGCAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGAT 617
          || ||||| |||||    ||| | || ||||||| |||   |||||| |  ||||||  |
     1103 CTAGAGTGTAGTGGTGCAATCACAGCTCACTGCAGCCTTGACCTCCCAGGCTCAAGCAGT 1162

      618 TCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCC----A 673
           |||   | |||||||||||||||||||||| |||||   ||  ||||||  |||    |
     1163 CCTCTCACTTCAGCCTCCCAAGTAGCTGGGACTACAGACACATACCACCATGCCCGTCTA 1222

      674 GCTTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAGTGGCAT 733
            ||||||| |||| |||||||||||| || ||||| ||||||||||||  ||||||
     1223 ATTTTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCGCCCAGGCTGGAGAGCAGTGGTGC 1282

      734 GATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTCAGTCTC 793
          ||||||  | |||||||  | |||   | |||||||||  | ||  ||||||||||||||| ||
     1283 GATCTCGGCTCACTGCAAGCTCTGCTTCCCGGGTTTATGCCATTCTCCTGCCTCAGCCTA 1342

      794 CTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA--TTTTTGTATTTTTA 851
          | ||||||||  ||| ||| |  |    || ||| |    |||||  |||||||| ||||
     1343 CCGAGTAGCTGGGACTACGGGTGCCTGCCACCATGCCTGGCTAATTTTTTTGTATCTTTA 1402

      852 GTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTCAGGTGAT 911
          |||||||  |  | |||   || |||||||||  |  | |||| ||  |  ||  |||||    |||||
     1403 GTAGAGACGGGGTTTCACCGTGTTGGCCAGGATGGTCTCGATCTCCTGACCT--TGTGAT 1460

      912 CCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGCCAGGCCT 971
          | ||  |||| || |||||||||||||||||||| |||| |||||||| | |||| |||||
     1461 CTGCCCGCCTCGGGCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACGCCTGGCCT 1520

      972 GAGCTACTCCTTTAGTCTCTGGAAAGACTGCGGCT 1006
            |  |    |||  | |  ||   | |   |||
     1521 AATTTTTGTATTTTTTGTAGAGATGAAGTCTTGCT 1555
```

.T 15
)-517-441-94/c
quence 94, Application US/10517441
)lication No. US20060121467A1
IERAL INFORMATION:
)PLICANT: FOEKENS, John
)PLICANT: HARBECK, Nadia
)PLICANT: KOENIG, Thomas
)PLICANT: MAIER, Sabine
)PLICANT: MARTENS, John
)PLICANT: MODEL, Fabian
)PLICANT: NIMMRICH, Inko
)PLICANT: RUJAN, Tamas
)PLICANT: SCHMITT, Armin
)PLICANT: SCHMITT, Manfred
)PLICANT: LOOK, Maxime P.
)PLICANT: MARX, Almuth
)PLICANT: HOEFLER, Heinz
:TLE OF INVENTION: Method and nucleic acids for the improved treatment of breast cel
:TLE OF INVENTION: proliferative disorders
:LE REFERENCE: 47675-93
)RRENT APPLICATION NUMBER: US/10/517,441
)RRENT FILING DATE: 2004-12-11
)IOR APPLICATION NUMBER: PCT/EP2003/010881
)IOR FILING DATE: 2003-10-01
)IOR APPLICATION NUMBER: DE 10317955.0

; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: DE 10300096.8
; PRIOR FILING DATE: 2003-01-07
; PRIOR APPLICATION NUMBER: DE 10245779.4
; PRIOR FILING DATE: 2002-10-01
; NUMBER OF SEQ ID NOS: 2147
; SEQ ID NO 94
; LENGTH: 2425
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (289, 833, 1773, 2369, 2374, 2386)
; OTHER INFORMATION: unknown base
US-10-517-441-94

Query Match 13.7%; Score 257; DB 6; Length 2425;
Best Local Similarity 75.5%; Pred. No. 3.1e-32;
Matches 367; Conservative 0; Mismatches 96; Indels 23; Gaps 3;

```
Qy        506 TTTTTTTTCTTTTTCTTTTTTTTTGAGACAGAGTCTTGCTCTGTCGCCCAGGCTGGAGTG 565
              |||   | | |    | |||||||||||||| ||||| |||| ||| ||||||||||||||
Db        635 TTTGGATACATGCCTTCTTTTTTTGAGACACAGTCTCGCTCAGTCACCCAGGCTGGAGTG 576

Qy        566 CAGTGGCATGATCTCTGCCCACTGCAACCTCTGCCTCCCGGATTCAAGCGATTCTCCTGC 625
              |||||||||||||||| || |||||||||||||| |||||||| |||||| ||||||||||
Db        575 CAGTGGCATGATCTCGGCTCACTGCAACCTCTACCTCCCGGGTTCAAGTGATTCTCCTGC 516

Qy        626 CTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCACGCCACCACACCCAGC---------- 675
              |||||||||||  ||| ||||||||||||||   |||   | |     |  |
Db        515 CTCAGCCTCCCGAGTGGCTGGGATTACAGGCACACTGTAGCTGGTACTACATGCCTGGCT 456

Qy        676 --------TTTTTTTATTTTGGAGACAGAGTCTTGCCCTGTCACCCAGGCTGGAGTACAG 727
                      |||||||  ||| |||||||||||| || || || ||||||||||||| |||
Db        455 AATTTTTTTTTTTTTTCTTTTGAGACAGAGTCTCGCTCTATCGCCCAGGCTGGAGTGCAG 396

Qy        728 TGGCATGATCTCAGTTCACTGCGACCTCCACCTCCCGGGTTCAAGCAATTCTCCTGCCTC 787
              ||||  |||||  | |||||||  |||||  | |||  |||||||  || |||||||  |||||
Db        395 TGGCGCGATCTGGGCTCACTGCAACCTCTTCTTCCAGGGTTCACGCCATTCTCCGGCCTC 336

Qy        788 AGTCTCCTGAGTAGCTAGGATTACAGAAGTGCACCTCCACGTTCAGCTAA---TTTTTGT 844
              ||  |||||||||||||  ||| ||||   |  |  |||  |  |  |  |   |||||||
Db        335 AGCCTCCTGAGTAGCTGGGACTACATGCGCCCGCAACCATGCCCGGNCTAATTTTTTTGT 276

Qy        845 ATTTTTAGTAGAGATGCGCTTTTGCCATGTTGGCCATGCTAGTCTGGAACCCCGGACCTC 904
              |||||||||||||||| | |||  |  |||| |||| | | ||||  ||| || ||||||
Db        275 ATTTTTAGTAGAGATGGGGTTTCACTGTGTTAGCCAGGATGGTCTCCAACTCCTGACCTC 216

Qy        905 AGGTGATCCGCTGGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCATCGCGC 964
              |  ||||||||   |||||||||||||||||||||||||| || |||||  |||||||   ||||
Db        215 A--TGATCCGCCCCCCTTGGCCTCCCAAAGTGCTGGGTTTACAGGCATGAGCCACTGCGC 158

Qy        965 CAGGCC 970
              | ||||
Db        157 CTGGCC 152
```

Search completed: June 21, 2006, 23:08:23
Job time : 63 secs